ELECTROSTATIC PROPERTIES OF PHASE-SEPARATING BOVINE LENS PROTEINS

by

Ian Shand Kovach

B.S. Physics, MIT

(1986)

Submitted to the Department of
Physics in Partial Fulfillment of
the Requirements for the
Degree of

DOCTOR OF PHILOSOPHY IN PHYSICS

at the

Massachusetts Institute of Technology

May, 1992

Signature of Author _____
Department of Physics
May 1992

Certified by _____
George B. Benedek
Alfred H. Caspary Professor of
Physics and Biological Physics
Thesis Supervisor

Accepted by _____
George Koster, Chairman
Departmental Committee
Department of Physics

ELECTROSTATIC PROPERTIES OF PHASE-SEPARATING BOVINE LENS PROTEINS

by

Ian Shand Kovach

Submitted to the Department of Physics in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Physics
May, 1992

ABSTRACT

Acid-base titration experiments were conducted to determine the net charge on four bovine lens proteins: γ_{II}, γ_{IIIa}, γ_{IIIb}, γ_{IV}. In addition to the dependence of protein charge on *pH*, the effects of ionic strength and identity on protein charge were investigated. Titration curves were obtained in 0.1 M KCl, 0.1 M NaCl, 0.1 M NaBr, and 0.01 M KCl.

Three theoretical methods, the Linderstrom-Lang model, a modification of the Linderstrom-Lang model which does not linearize the Poisson-Boltzmann equation, and the Kirkwood-Tanford model were used to correct intrinsic proton binding energies in an effort to predict protein charge as a function of solution conditions. The electrostatic interaction energy between protein and solvent was determined using the Linderstrom-Lang model.

The acid-base titration curves of the four proteins were found to be very similar, with the exception of basic range titration of γ_{IIIb} which exhibited slightly greater negative charge near *pH* 10. No evidence of protein charge dependence on the identity of electrolyte was observed. A dependence of protein charge on ionic strength was observed. Decreasing ionic strength was found to correlate with decreasing magnitude of net protein charge at a given *pH*.

The theoretical Linderstrom-Lang titration curves were found to be in fair agreement with the experimental results over a *pH* range of 5 to 11, and in poor agreement in the *pH* extremes. The nonlinearized variation of the Linderstrom-Lang model was found to be in good agreement with the titration curves from *pH* 2 to *pH* 11. The theoretical Kirkwood-Tanford titration curves were determined for γ_{II} and also found to be in excellent agreement with experimental results. The precise location of protein charge on γ_{II}, as a function of *pH*, was determined using the Kirkwood-Tanford model.

Thesis Supervisor: George B. Benedek
Title: Alfred H. Caspary Professor of Physics and Biological Physics

1 - Introduction

This work characterized the charge distribution of a family of phase-separating bovine lens proteins, the γ-crystallins. The net protein charge as a function of *pH* was experimentally determined by means of acid-base titration. Electrostatic interactions within the proteins and between the proteins and solvent were studied. These calculations predicted the exact location of charge as a function of *pH* on γ_{II}. They also predict the work required to charge these proteins and the electrostatic interaction energy between the protein and solvent.

Chapter 2 discusses the theory of proton binding. A detailed description of proton binding is presented in terms of the grand canonical ensemble. Binding at both single and multiple sites is considered. The results of this treatment are compared to the classical description of Henderson and Hasselbalch to provide a relationship between *pK*, proton chemical potential in solution, and the proton binding energy. This expression allows for the quantitative description of the effects of intraprotein electrostatic interactions on *pK*s.

Chapter 3 is a detailed account of the acid-base

titration experiments. A review of the physics of *pH* determination is provided. The design of the glass membrane *pH* electrode is explained and all of the contributions to the voltage it measures are discussed. Experimental uncertainties are discussed in this chapter.

In chapter 4, the conversion of raw data to the form charge versus *pH* is discussed. Two techniques to accomplish this are presented, and the need for control experiments with water is explained. The propagation of experimental uncertainties is discussed. The protein titration data is presented in this chapter. In addition to studying the charge as it depends on *pH*, the effects of other solvent conditions, such as ionic strength and electrolyte identity, were investigated and are discussed here.

Three theories with which to describe intraprotein electrostatic interactions are described in chapter 5. These include the Linderstrom-Lang model and the Kirkwood-Tanford model. In addition, a modification of the Linderstrom-Lang model in which the Poisson-Boltzmann equation is not linearized, is presented. This is, to the best of my knowledge, the first time that the nonlinearized Poisson-Boltzmann equation has been used to predict the titration behavior of a charge shell model protein.

The three methods are used to study the effects of electrostatic interactions on proton binding. The work associated with charging the proteins is also presented in this chapter as is the protein-solvent interaction energy.

The relationship between protein solvation energy and the critical temperature is briefly discussed. The Kirkwood-Tanford model requires detailed structural information, including solvent accessibility. The necessary information was only available for γ_{II}, and as a result some of this chapter deals only with γ_{II}.

Chapter 6 concisely recounts the findings of the earlier chapters. Some possible directions for related work in the future are also discussed.

This work is the first attempt to comprehensively investigate the electrostatic properties of the γ-crystallins. These proteins, and their homologues in other species, play a key role in the pathogenesis of cataracts. For this reason alone, as much information as possible must be made available about these proteins in order to expedite improved therapies for this significant public health problem. This work also constitutes an important step in the development of a complete theory of liquid-liquid phase separation in protein solutions. Understanding the charge on the γ-crystallins, and the associated protein-solvent interaction energy, is a prerequisite for future work aimed at relating intermolecular interactions to the location of the phase boundaries.

Chapter 2 - Theory of Proton Binding

2.0 Introductory Remarks

This chapter reviews the Grand Canonical Ensemble to the extent necessary to describe proton binding to proteins. It then turns to a derivation of the proton's chemical potential in solution and its relationship to *pH* and activity. Specific aspects of proton binding are discussed, beginning with binding at one site, and moving on to binding at multiple sites and the role of electrostatic interactions in altering proton binding energies. Finally, fluctuation in proton binding is discussed.

The Grand Canonical Ensemble, and its application to the general problem of binding, is discussed in most Statistical Mechanics texts[1,2]. The material in this chapter specific to proton binding parallels descriptions in the literature[3]. The approach taken here, however, differs somewhat.

---

[1]Landau, L.; Lifshitz, E. *Statistical Physics;* Pergamon: Oxford, 1980.
[2]Reif,F. *Fundamentals of Statistical and Thermal Physics;* McGraw-Hill: New York, 1965.
[3]Tanford, C. *The Physical Chemistry of Macromolecules;* Wiley: New York, 1961.

2.1 The Grand Canonical Ensemble

In the field of Statistical Mechanics, the Grand Canonical Ensemble provides a formal method for describing thermodynamic systems in which both particles and energy can be exchanged. This approach makes reference to a larger isolated system in which energy and particle number are constant. Those systems which are free to exchange particles and energy are considered to be contained within the larger system and will be referred to as subsystems. Since the subsystems are taken to be much smaller than the isolated system, the isolated system can be thought of as a reservoir of particles and energy.

The total energy and particle number in the large system, denoted with the subscript o, can be separated into components within a particular subsystem and components in the remaining reservoir. These are denoted by the subscripts i and r respectively, allowing the conservation of energy and particle number to be written as,

$$E_o = E_i + E_r \quad , \tag{2.1}$$

$$N_o = N_i + N_r \quad . \tag{2.2}$$

The probability of finding a given E_i, N_i combination within a subsystem can be written in terms of the number of

thermodynamic states available to the reservoir Ω_r, and to the subsystem Ω_i, at their respective energies and particle numbers. This probability, denoted by $P_i(E_i,N_i)$, can be written as,

$$P_i(E_i,N_i) = \alpha' \, \Omega_r(E_o-E_i,N_o-N_i) \, \Omega_i(E_i,N_i), \tag{2.3}$$

where α' is a proportionality constant. Absorbing $\Omega_r(E_o,N_o)$, which is a constant, into the proportionality constant allows equation 2.3 to be rewritten as,

$$P_i(E_i,N_i) = \alpha \, \frac{\Omega_r(E_o-E_i,N_o-N_i)\,\Omega_i(E_i,N_i)}{\Omega_r(E_o,N_o)} \tag{2.4}$$

or,

$$P_i(E_i,N_i) = \alpha\Omega_i(E_i,N_i) \exp\left([S_r(E_o-E_i,N_o-N_i) - S_r(E_o,N_o)]/k \right) \tag{2.5}$$

having used the definition of the entropy in terms of the number of thermodynamic states, $\Omega = e^{S/k}$, where S is entropy.

$S(E_o-E_i,N_o-N_i)$ can be expanded around the point E_o, N_o, noting that $E_i << E_o$ and $N_i << N_o$, i.e.,

$$S_r(E_o-E_i,N_o-N_i) \approx S_r(E_o,N_o) - \left(\frac{\partial S_r}{\partial E}\right)(E_i) - \left(\frac{\partial S_r}{\partial N}\right)(N_i) \tag{2.6}$$

where the derivatives are evaluated at (E_o,N_o). Recalling

that,

$$\frac{\partial S}{\partial N} = \frac{-\mu}{kT} \quad \text{and} \quad \frac{\partial S}{\partial E} = \frac{1}{kT} \tag{2.7}$$

where μ is the chemical potential of the exchangeable particles, and T is temperature.[4] The probability of finding the combination of particle number and energy N_i and E_i respectively in the subsystem, can be rewritten as,

$$P_i = \alpha\ \Omega_i\ exp(-\beta E_i + \beta\mu N_i) \tag{2.8}$$

where $\beta = \frac{1}{kT}$.

The proportionality constant α is identified as one over the grand canonical partition function $\mathcal{Z}$, and can be determined by the normalization requirement that $\sum P_i(E_i, N_i) = 1$, i.e.,

$$\mathcal{Z} = \frac{1}{\alpha} = \sum_N Z_N(\beta, V)\ e^{\beta\mu N} \ . \tag{2.9}$$

In equation 2.9, Z_N is the canonical partition function determined at a given N, and can be written as,

$$Z_N = \sum_j e^{-\beta E_j} \ . \tag{2.10}$$

---

[4]Zemansky, M. W. *Heat and Thermodynamics;* McGraw-Hill: New York, 1957.

The sum is over all possible configurations of the subsystem at fixed number N.

By first calculating the grand partition function, several important features of a subsystem can in turn be calculated with relative ease, e.g., the number of particles, the entropy, the Gibbs energy, etc. Determining the number of particles in a subsystem is of special interest in the study of proton binding to proteins. Therefore, the relationship between the grand partition function and the number of particles in a subsystem will be reviewed in detail.

The average number of particles in a subsystem can be written as,

$$\langle N \rangle = \sum_{N=0}^{\infty} N\, P(N) \qquad (2.11)$$

where $P(N)$ is the probability of finding N particles in the subsystem. $P(N)$ can in turn be written in terms of the chemical potential μ , and the canonical partition function Z_N, as,

$$P(N) = \frac{e^{\beta\mu N}\, Z_N}{\mathcal{Z}} \qquad (2.12)$$

with reference to equations 2.8 and 2.9. This allows $\langle N \rangle$ to

be written as,

$$\langle N \rangle = \frac{\sum N e^{\beta\mu N} Z_N}{\sum e^{\beta\mu N} Z_N} = kT \frac{\partial}{\partial \mu} \ln \mathcal{Z}. \qquad (2.13)$$

By following a similar argument, the fluctuation in N can be shown to be,

$$\langle (N - \langle N \rangle)^2 \rangle = kT \frac{\partial}{\partial \mu} \langle N \rangle. \qquad (2.14)$$

The Grand Canonical Ensemble provides a powerful technique for characterizing systems in which particles and energy can be exchanged. In order to apply this technique to proton binding, protein molecules or individual binding sites will be treated as subsystems, and the solution in which protein molecules are dissolved as the reservoir. Before implementing this technique, however, an understanding is needed of the proton chemical potential in solution and of the proton binding energies to protein.

2.2 Protons in Solution: Chemical Potential, *pH*, and Activity

In section 2.1 the number of particles found in a thermodynamic system which is free to exchange particles was

cast in terms of the particles' chemical potential. In order to apply the formalism of section 2.1 to a specific case, e.g., proton binding, an explicit expression for the chemical potential is required. In this section, the chemical potential of protons in solution is derived using a lattice model. The chemical potential is then related to the more commonly used descriptive parameters, *pH* and activity.

The chemical potential of a particle is the change in the free energy that arises from the addition of one particle. Another way of expressing this is as the partial derivative of the free energy with respect to the number of particles. In the case of protons,

$$\mu_{H^+} = \frac{\partial G}{\partial N_{H^+}} \tag{2.15}$$

where G is the Gibbs free energy of the aqueous solvent in which unbound protons are found. The free energy can be broken into components, namely the enthalpy and the entropy of mixing,

$$G = H - TS \tag{2.16}$$

where H is the enthalpy, T the temperature, and S the entropy. H includes the self energies of all of the constituent particles as well as all interaction energies.

Proceeding further requires an expression for the entropy in terms of the number of protons in solution and the solution volume. In order to obtain this, the fluid is treated as a lattice in which sites are to be filled by either protons or water molecules. Denoting the total number of lattice sites as N_o and the total number of protons as N, the number of possible configurations of the solution can be written as,

$$\Omega = \frac{N_o!}{N!\ (N_o-N)!}\ . \tag{2.17}$$

Noting the definition of the thermodynamic entropy as $S = k ln\Omega$, and making use of Stirling's approximation, i.e. $ln(N!) = Nln(N) - N$ for large N, the entropy of the solution can be written as,

$$S \approx - Nk\ ln\left(\frac{N}{N_o-N}\right) + N_o k\ ln\left(\frac{N_o}{N_o-N}\right). \tag{2.18}$$

Imposing the assumption that $N<<N_o$,

$$\frac{N}{N_o-N} \approx \frac{N}{N_o} \quad \text{and} \quad \frac{N_o}{N_o-N} \approx 1 \tag{2.19}$$

leaving,

$$S \approx - Nk\ ln\left(\frac{N}{N_o}\right) \tag{2.20}$$

The legitimacy of the dilute assumption, $N<<N_o$, can be seen by noting that the molarity of water molecules in pure water is 55 M, while the molarity of hydrogen ions in aqueous solution ranges between at 10^{-2} M and 10^{-12} M for pH between 2 and 12.

Returning to equations 2.15 and 2.16, the chemical potential is written as,

$$\mu_{H^+} = \frac{\partial H}{\partial N_{H^+}} - T \frac{\partial S}{\partial N_{H^+}} \qquad (2.21)$$

or, making use of the model entropy of equation 2.20,

$$\mu_{H^+} = \mu_o(P,T) + kT \ln(\chi_H) \qquad (2.22)$$

where χ_H is the number fraction of protons in solution, P is pressure, and T is temperature. The "standard" component of the chemical potential, μ_o, represents the enthalpic change associated with the addition of one proton to solution. In the dilute limit the enthalpic contribution to the chemical potential remains concentration independent. This is the case since effectively all interactions of the added proton are with water molecules. When the dilute limit is violated there is a concentration dependence of the enthalpy change resulting from interactions between protons and between protons and their counterions. This is denoted by $-h(\chi_H)$.

The chemical potential then takes the form,

$$\mu_{H^+} = \mu_o(P,T) + kT\ ln(\chi_H) - h(\chi_H) \quad . \tag{2.23}$$

This new term can be incorporated into the logarithm to give,

$$\mu_{H^+} = \mu_o(P,T) + kT\ ln(\chi_H\ e^{-\beta h(\chi_H)}) \tag{2.24}$$

where again, $\beta = 1/kT$. The exponentiated factor is defined as the activity coefficient, f. By convention, the proton concentration is usually written in terms of moles per liter as opposed to number fraction. Also by convention, the base ten logarithm is often used. Equation 2.24 then becomes,

$$\frac{\mu_{H^+} - \mu_o}{kT\ ln(10)} = log_{10}(f[H^+]) \quad . \tag{2.25}$$

The log term on the right is defined as negative the pH. The activity, α, is defined as the concentration of proton multiplied by the activity coefficient, i.e. $\alpha = f[H^+]$.

2.3 Proton Binding at One Site

By utilizing the chemical potential of protons in solution from section 2.2, the formalism of section 2.1 can

be used to describe proton binding. A single binding site is considered a subsystem, as described in section 2.1, and the solution a reservoir of both protons and energy. Recognizing that the occupation number of a single site only can be zero or one, the grand partition function takes on the simple form,

$$\mathcal{Z} = z_0 + z_1 e^{\beta\mu} \tag{2.26}$$

where z_0 and z_1 are the canonical partition functions associated with the $N = 0$ and $N = 1$ states respectively. These canonical partition functions can be evaluated using equation 2.10,

$$z_0 = 1 \tag{2.27}$$

$$z_1 = e^{-\beta\varepsilon} \tag{2.28}$$

where ε is the proton binding energy of the site. This allows the grand partition function to be written as,

$$\mathcal{Z} = 1 + e^{\beta(\mu-\varepsilon)}. \tag{2.29}$$

The average occupancy can then be determined by making use of equation 2.13, which gives,

$$\langle N \rangle = \frac{1}{1 + e^{\beta(\varepsilon-\mu)}} \cdot \qquad (2.30)$$

This familiar result represents the probability of occupancy for any single state system obeying Fermi-Dirac statistics. In the case of proton binding, the energy, ε, is the proton binding energy. A detailed discussion of the chemical residues responsible for proton binding in proteins and their respective intrinsic binding energies is presented in section 2.6.

2.4 The Dissociation Constant and the Henderson-Hasselbalch Equation

The more standard description of proton binding starts from a different perspective where the notion of a dissociation constant replaces binding energy. The dissociation constant is defined as,

$$K = \frac{[H^+][A^-]}{[HA]} \qquad (2.31)$$

where the brackets indicate the molar concentration of protons in solution, dissociated binding sites, and bound protons, respectively. By taking the logarithm of this equation and making the definitions $pK = -log_{10}(K)$ and

tentatively $pH = -log_{10}([H^+])$, where the activity coefficient is taken to be one (see section 2.4), equation 2.31 can be rewritten as the classic Henderson-Hasselbalch equation,

$$pH = pK + log_{10}\frac{[A^-]}{[HA]} \quad . \tag{2.32}$$

Note that $\langle N \rangle$ is expressed in terms of concentrations as,

$$\langle N \rangle = \frac{[HA]}{[A^-] + [HA]} \tag{2.33}$$

or, making use of equation 2.31,

$$\langle N \rangle = \frac{1}{1 + K/[H^+]} \quad . \tag{2.34}$$

Given the definition of pH and pK, equation 2.34 can be rewritten as,

$$\langle N \rangle = \frac{1}{1 + 10^{pH-pK}} \quad . \tag{2.35}$$

Figure 2.1 compares this theoretical curve with experimental titration data obtained for acetate which exhibits one proton binding site per molecule. The manner in which the experimental titration data was collected will be discussed in chapter 3. As figure 2.1 demonstrates, the theoretical

titration curve is in excellent agreement with experiment.

Equation 2.35 must be consistent with equation 2.30. Section 2.5 compares these descriptions in order to determine the relationship between the dissociation constant and the binding energy.

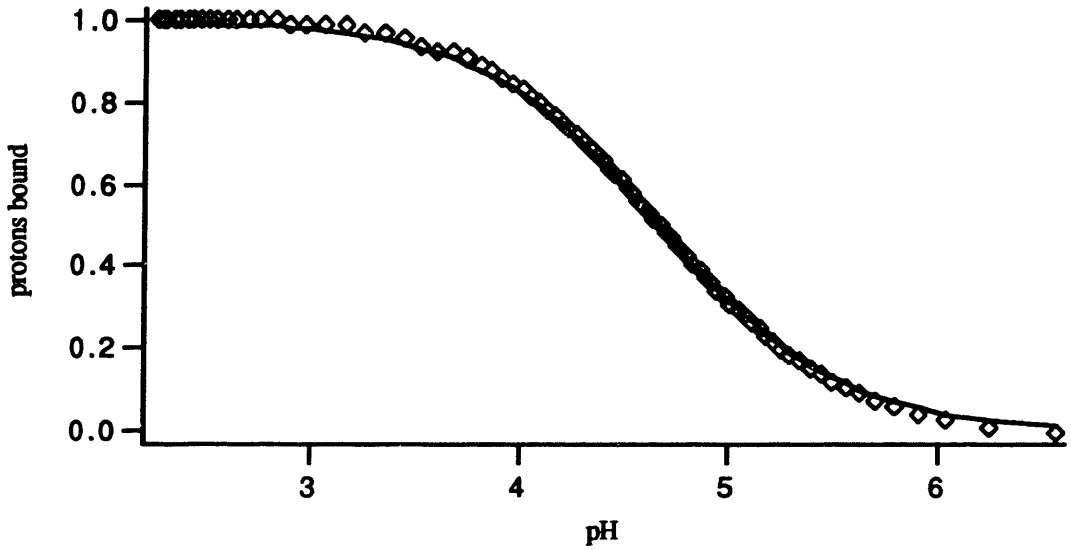


Figure 2.1 - Diamonds indicate titration data for acetic acid which has one titrateable site. This data was collected in a manner described in chapter 3. The solid line represents the theoretical curve obtained by using a *pK* of 4.67 in equation 2.35.

2.5 The Relationship Between Proton Binding Energy and *pK*

This section examines the relationship between the binding energy of the proton and the *pK*. A working knowledge

of this relationship will allow quantitative predictions regarding the effect on *pK*s of intramolecular electrostatic interactions which alter the proton binding energy. Comparing equation 2.30 to equation 2.35 reveals the following relationship,

$$10^{pH-pK} = e^{\beta(\varepsilon-\mu)}. \tag{2.36}$$

Using the form of the chemical potential in equation 2.22 allows further clarification of the significance of *pK*, specifically,

$$10^{-pK} = e^{\beta(\varepsilon-\mu o)} \tag{2.37}$$

or,

$$pK = \frac{\mu_o - \varepsilon}{kT\ ln(10)}\ . \tag{2.38}$$

This equation relates the *pK* to the binding energy and the standard part of the chemical potential. Section 2.8 will make use of this to determine the effect electrostatic interactions have on *pK*s. Before turning to that topic it is necessary to discuss the intrinsic proton binding energies of the titrateable residues in the absence of extrinsic interactions.

2.6 Intrinsic *pKs* of the Amino Acids in Protein

The chemical groups which participate in proton binding to proteins are portions of amino acids. The types of amino acids involved in proton binding fall broadly into two classes: "acidic" and "basic" (see figure 2.2). Acidic residues are those that become negatively charged with the loss of a proton. Basic residues are those that become neutral, carrying positive charge when a proton is bound. Generally, acidic residues exhibit loss of their proton in the acidic range of *pH* (below 7) while basic residues exhibit loss of their proton in the basic *pH* range (above 7).

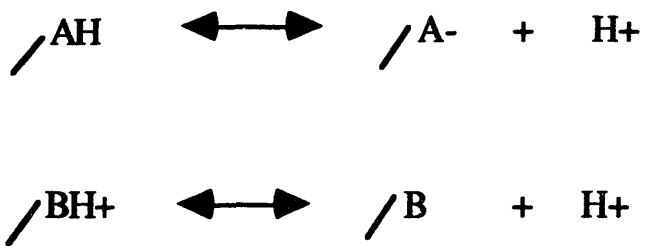


Figure 2.2 - Proton binding reaction at acidic versus basic residues.

Equation 2.35 can be used to introduce the concept of intrinsic *pKs* for the amino acids. Intrinsic *pK* refers to the *pK* in the absence of any external electrostatic

interactions. These quantities are required for theoretical predictions of *pK*s which will be discussed in chapter 5. The intrinsic *pK*s and the corresponding difference between binding energy and standard chemical potential are listed in table 2.1 for titrateable amino acids.[5]

Table 2.1

Residue	*pK*	$(\mu_o - \varepsilon)(kT)$	class
R-chain			
Aspartic Acid	4.0	9.2	acidic
Glutamic Acid	4.5	10.4	acidic
Histidine	6.4	14.7	basic
Tyrosine	10.0	23.0	acidic
Lysine	10.4	23.9	basic
Arginine	12.0	27.6	basic
Terminal Carboxylic	3.6	8.3	acidic
Terminal Amino	10.4	23.9	basic

These *pK*s do not correspond to the *pK*s observed for amino acids free in solution. This is because amino acids free in solution are capable of carrying charge at α-carboxyl and α-amino sites which can alter the observed *pK*s. In proteins, however, these sites are involved in peptide bonds and are hence uncharged. Note that the α-amino and α-

[5]Tanford, C. *Adv. Prot. Chem.* **1962**, 17, 69.

carboxyl groups of amino acids at the first and last residue of a peptide, respectively, are ionizable. The values of the *pK*s given in table 2.1 correspond to the *pK*s observed for amino acids in denatured protein.[6] These values are chosen because in denatured protein, charge interactions are minimized by virtue of increased intercharge separation resulting from unfolding of the protein.

2.7 Proton Binding at Multiple Sites Without Interactions

This section discusses the problem of multiple proton binding sites in the absence of extrinsic charge interactions, making use of the formalism presented in section 2.1. In this context a protein molecule with multiple binding sites represents the subsystem described in section 2.1.

Consider a protein with M different classes of ionizable amino acid residues, each type denoted by the subscript k. Each class is taken to have a different proton binding energy, ε_k. Assume there are N_k members of each class, and denote the number of protons bound to a class with ν_k, where $0 \leq \nu_k \leq N_k$. Using this notation, the total energy associated with proton binding to a given protein can be written as,

[6]Tanford, C. *Adv. Prot. Chem.* **1962**, 17, 69.

$$\varepsilon(\{\nu_k\}) = \sum_{k=1}^{M} \varepsilon_k \nu_k \tag{2.39}$$

The total number of protons bound to a protein is written as ν, where,

$$\nu = \sum_{k=1}^{M} \nu_k \tag{2.40}$$

and the total number of proton binding sites is denoted by N, where,

$$N = \sum_{k=1}^{M} N_k \tag{2.41}$$

Referring to section 2.1, the canonical partition function can be written in terms of these parameters,

$$\mathcal{Z}(T, V, \nu) = \sum_{\nu_N} \cdots \sum\sum_{\nu_2 \nu_1} g(\{\nu_k\})\ e^{-\beta \Sigma \varepsilon_k \nu_k} \tag{2.42}$$

where g is a degeneracy factor which depends on the particular sequence $\{\nu_k\}$. For a given set of ν_ks the degeneracy factor, g, can be written as,

$$g(\{\nu_k\}) = \binom{N_M}{\nu_M} \cdots \binom{N_2}{\nu_2}\binom{N_1}{\nu_1} = \Pi \binom{N_k}{\nu_k} \tag{2.43}$$

allowing the canonical partition function to be rewritten as,

$$\mathcal{Z}(T, V, \nu) = \sum_{\nu_N} \cdots \sum_{\nu_2 \nu_1} \sum \prod \binom{N_k}{\nu_k} e^{-\beta \Sigma \varepsilon_k \nu_k} \tag{2.44}$$

subject to the constraint that $\Sigma \nu_k = \nu$. This expression can then be used with equation 2.9 to determine the grand canonical partition function,

$$\mathcal{Z}(T, V, \mu) = \sum_{\nu_N}^{N_M} \cdots \sum_{\nu_2 \nu_1}^{N_2 N_1} \sum \prod \binom{N_k}{\nu_k} e^{\beta(\mu - \varepsilon_k)\nu_k} \tag{2.45}$$

or,

$$\mathcal{Z}(T, V, \mu) = \sum_{\nu_N}^{N_M} \binom{N_M}{\nu_M} e^{\beta(\mu-\varepsilon_M)\nu_M} \cdots \sum_{\nu_2}^{N_2} \binom{N_2}{\nu_2} e^{\beta(\mu-\varepsilon_2)\nu_2} \sum_{\nu_1}^{N_1} \binom{N_1}{\nu_1} e^{\beta(\mu-\varepsilon_1)\nu} \tag{2.46}$$

or, finally,

$$\mathcal{Z}(T, V, \mu) = \left(1+e^{\beta(\mu-\varepsilon_M)}\right)^{N_M} \cdots \left(1+e^{\beta(\mu-\varepsilon_2)}\right)^{N_2} \left(1+e^{\beta(\mu-\varepsilon_1)}\right)^{N_1}. \tag{2.47}$$

This in turn allows the determination of the proton binding curve as a function of *pH*, making use of equations 2.13 and 2.36,

$$\langle N \rangle = \sum_{k=1}^{M} \frac{N_k}{1 + 10^{pH - pK_k}}. \tag{2.48}$$

As expected in the absence of interactions, this equation is

the sum of the individual binding site curves.

2.8 Proton Binding in the Presence of Electrostatic Interactions

The description of proton binding in the presence of electrostatic interactions can be understood in the context of equation 2.38. Using the notion of intrinsic *pK*s and binding energies as discussed in section 2.6, the *pK* at a particular proton binding site can be written in terms of the site's intrinsic binding energy and an electrostatic correction,

$$pK = \frac{\mu_o - (\varepsilon_o + W)}{kT\ ln(10)} \qquad (2.49)$$

where ε_o is the binding energy in the absence of electrostatic interaction and W is the change in the binding energy which arises by virtue of interactions with other sites. This can be rewritten in terms of the intrinsic *pK* as,

$$pK = pK_o + \Delta pK \qquad (2.50)$$

where,

$$\Delta pK = \frac{-W}{kT\ ln(10)}. \qquad (2.51)$$

By virtue of its pair-wise additive nature, the electrostatic interaction energy can be broken into terms representing the contributions from individual charged sites within the protein, indexed by j. Considering site i, we can write,

$$W_i = \sum_{j \neq i} W_{ij}. \qquad (2.52)$$

Here, the W_{ij} indicates the electrostatic interaction between ionizable sites i and j. In principle this can be very difficult to evaluate. Understanding these interactions forms a significant portion of chapter 5, where ideal model protein structures will be used to calculate the pK changes in the bovine lens protein γ_{II}.

2.9 Charge Fluctuation

As demonstrated in section 2.1, the fluctuation in the number of bound protons, which corresponds to the fluctuation in charge, is just kT times the derivative with respect to the chemical potential of the average number of protons. By

using equation 2.14 and the definition of *pH*, the charge fluctuation can be written in terms of *pH* as,

$$\langle \delta N^2 \rangle = - \frac{1}{2.303} \frac{\partial \langle N \rangle}{\partial pH}. \qquad (2.53)$$

The charge fluctuation is proportional to the slope of the proton binding curve as a function of *pH*. Applying this to the special case of one binding site reveals charge fluctuation of the form,

$$\langle \delta N^2 \rangle = \frac{10^{pH-pK}}{(10^{pH-pK} + 1)^2}. \qquad (2.54)$$

The charge fluctuation is of interest because of its role in interprotein interactions.[7]

[7]Kirkwood, J. G.; Shumaker, J. B. *Proc. Natl. Acad. Sci.* **1952**, 38, 863.

Chapter 3 - Acid-Base Titration Experiments

3.0 Introductory Remarks

This chapter describes acid-base titration experiments performed with the bovine lens proteins γ_{II}, γ_{IIIa}, γ_{IIIb}, and γ_{IV}. In these experiments acid or base was added to aqueous electrolyte solutions containing dissolved, purified protein of known concentration. The *pH* was monitored in a continuous fashion as the acid or base was added. The experiment was fully automated so that both the addition of acid or base and the measurement of *pH* were interfaced with a computer. The amount of acid or base added was monitored and was recorded with the measured *pH*. The resulting data files were of the form *pH* versus volume of acid or base added.

In order to monitor the *pH* in the solution, a glass membrane electrode was utilized. In this type of electrode, a thin glass membrane which is selectively permeable to hydrogen ions, separates the solution being studied from a reference solution which contains a metal electrode. The second metal electrode is in contact with the solution being

studied by means of a conducting solution - saturated KCl - which is referred to as a salt bridge. The electrodes are connected to a *pH* meter which is essentially a voltmeter, providing extremely high resistance. The *pH* dependent component of the potential difference arises from the glass membrane. In equilibrium, the chemical potential of the hydrogen ions must be equal on either side of the membrane. Since the concentrations of hydrogen ion are different across the membrane, a hydrogen ion current will flow until an electrostatic potential develops across the membrane, causing the chemical potentials to be equal on either side. The details of this will be discussed in section 3.2.2.

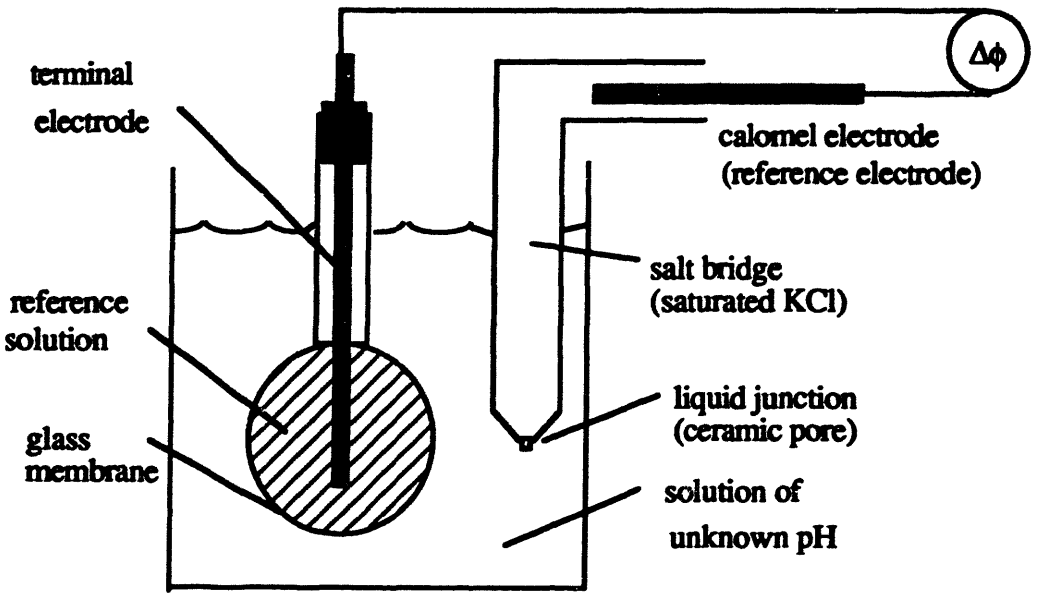


Figure 3.1 -- Schematic illustration of the glass membrane electrode adapted from Vetter[8]. See text for details

[8]Vetter, K. J. *Electrochemical Kinetics;* Academic Press: New York, 1967.

In addition to describing the *pH* electrode in more detail, this chapter provides an overview of the apparatus used and a description of the experimental procedure involved in these acid-base titration experiments. This includes a description of sample preparation and handling. Experimental uncertainties are discussed and representative examples of the raw data are presented.

3.1 Apparatus

The apparatus used in the acid-base titration experiments consisted of a Hamilton Microlab M automated dispenser, DEC MicroVAX II, Radiometer GK473902 *pH* electrode, Radiometer *PH* M85 *pH* meter, and glassware to hold the protein solution and isolate it from the atmosphere. The Hamilton Microlab M dispenser is capable of delivering variable aliquot sizes to below 5 microliters with accuracy to better than 1%. It was used to deposit aliquots of acid or base into a 20 ml flask containing the protein solution. The Radiometer GK473902 electrode is a glass membrane, mercury-mercurous chloride combination *pH* electrode. It was used to continuously monitor the *pH* in the protein solution as acid or base was added. The *pH* meter was connected via RS232 to the MicroVAX. The temperature of both the protein solution

and the acid or base used to titrate was maintained by submersion in water baths of both the flask containing the protein solution and the flask containing the acid or base. The temperature of the water baths was controlled by thermal contact with the reservoir of a Nesslab Endocal circulating bath. See figure 3.2.

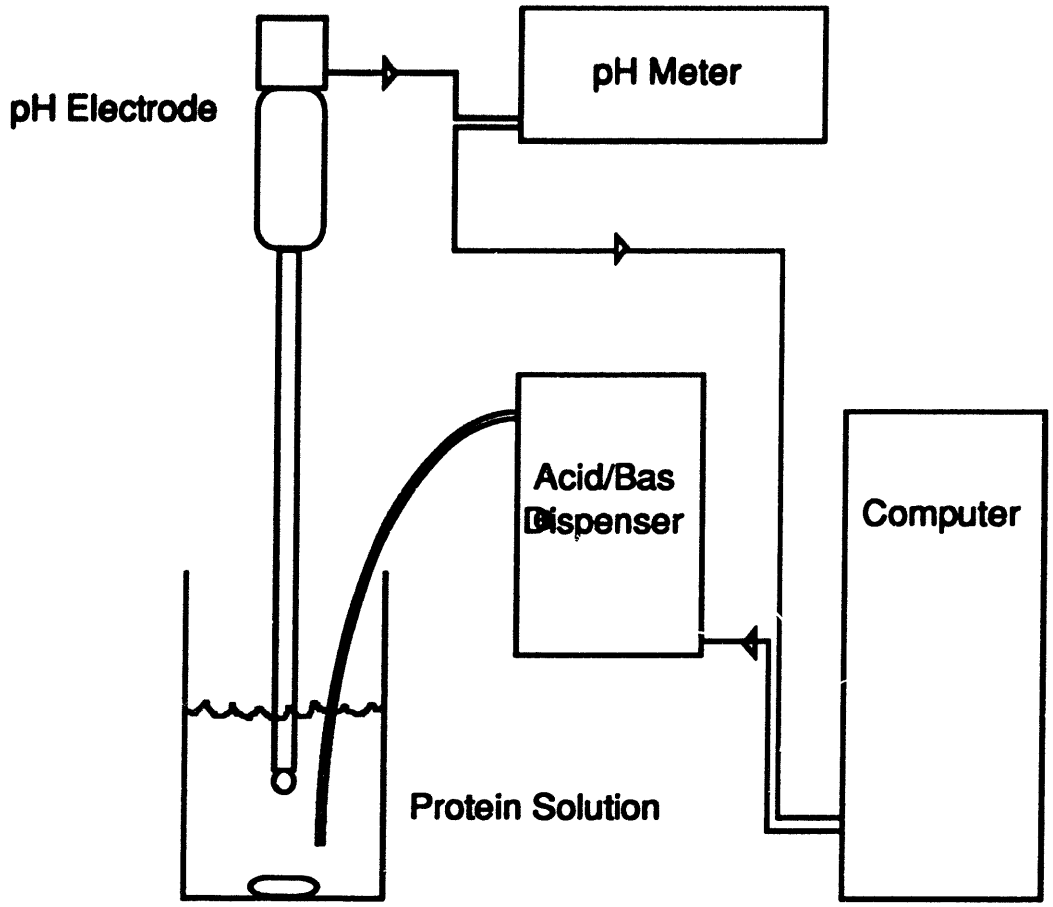


Figure 3.2 -- The acid-base titration apparatus.

The software which governed the acid/base dispenser was written by Michael Orkiscz, and is reproduced in appendix A.

The general procedure was for the MicroVax to evaluate the *pH* as read by the *pH* meter at specified intervals (e.g. every second). This process continued without activating the dispenser until the standard deviation of a specified number of sequential measurements, *N*, fell below a specified tolerance limit. At this point, the MicroVax recorded the mean of the last *N* measurements and activated the dispenser to dispense a specified volume of acid or base into the protein solution. Typical values used for *N* and tolerance respectively, were 5 and .01 *pH* units. Any sequence of aliquot sizes could be dispensed. Varying the aliquot size and choosing appropriate concentrations of acid or base rendered a uniformly high density of data over a large *pH* range.

3.2 The Combined Glass Membrane *pH* Electrode

To fully understand the combined glass membrane electrode it is necessary to consider its two components, the glass membrane electrode and the reference electrode, separately.[9,10,11] A schematic illustration of the combined glass membrane electrode is presented in figure 3.1.

---

[9]Bates, R. G. *J. Res. Natl. Bur. Std.* **1950**, 45, 418.
[10]Lakshminarayanaiah, N. *Membrane Electrodes;* Academic Press: New York, 1976.
[11]Eisenman, G. *Glass Electrodes for Hydrogen and Other Cations;* Marcel Dekker: New York, 1967.

The glass membrane electrode consists of a very thin-walled glass bulb (≈0.005 mm thick) which is permeable to hydrogen ions. The bulb is filled with a reference solution of known *pH* in which a metal terminal electrode is immersed. The reference solution in the bulb must have a stable *pH* and it must form a stable junction potential with the metal electrode in the bulb. The most common choices of reference solution for platinum probes are standard acetate (0.1 M acetic acid plus 0.1 M sodium acetate) or 0.1 N HCl. The thin-walled glass bulb is then placed in the solution of unknown *pH* and a potential difference develops across the glass membrane of the bulb. The source of this potential difference relates to the differing hydrogen ion concentrations on the two sides of the glass membrane. The relationship between this potential difference, $\Delta\phi_{glass}$, and the *pH* of the solution being investigated will be explored in detail in section 3.2. In order to measure this potential difference however, a reference electrode must also be placed in the solution.

The reference electrode used in these experiments consisted of calomel (mercurous, mercurous chloride) in contact with a salt bridge of saturated KCl. The saturated KCl was, in turn, in contact with the solution being investigated by means of a liquid junction. See figure 3.1. The ideal reference electrode should exhibit a constant, reproducible potential which is independent of the *pH* of the

solution being studied. The potentials arising from the calomel reference electrode are to good approximation constant and independent of solution *pH*, making the calomel electrode an excellent choice as a reference electrode for *pH* measurements. In order to make the conversion from potential difference to *pH*, the *pH* meter requires calibration with solutions of known *pH*.

The *pH* electrode which was used in our experiments is illustrated in detail in figure 3.3. The four interfaces which give rise to the potential difference this electrode measures are: (1) the glass membrane; (2) the junction between the platinum terminal electrode in compartment I and the reference solution; (3) the junction between saturated KCl and calomel in compartment II; and (4) the liquid junction potential across the ceramic pore which separates the solution being studied from the saturated KCl solution in compartment II. The net result is a potential difference which can be written as,

$$\Delta\phi = \Delta\phi_{glass}+\Delta\phi^{I}_{metal\text{-}liquid}+\Delta\phi^{II}_{metal\text{-}liquid}+\Delta\phi_{liquid\ junct} \qquad (3.1)$$

Each term will be considered separately, and the dependence of $\Delta\phi$ on the *pH* of the solution considered.

3.2.1 Liquid-Metal Junctions in Compartments I and II

Determining an exact expression for the liquid-metal junctions is difficult. Fortunately it is not necessary to do so in order to determine the *pH* of a solution into which the electrode has been placed. The liquids in both compartments I and II are unaffected by the *pH* of the solution being investigated. As a result, the junction potentials arising from the liquid-metal interfaces in these compartments are independent of the *pH*, of the solution being studied. These junction potentials are, to very good approximation, constant over the course of an experiment, i.e.,

$$\Delta\phi^{I}_{junct} + \Delta\phi^{II}_{junct} \approx C \qquad (3.2)$$

This leaves the total potential produced by the electrode as,

$$\Delta\phi \approx \Delta\phi_{glass} + \Delta\phi_{liquid\ junct} + C \qquad (3.3)$$

Because C is constant, it is irrelevant to the determination of *pH*. Because the electrode is calibrated with solutions of known *pH*, only changes of $\Delta\phi$ are involved in the determination of *pH*.

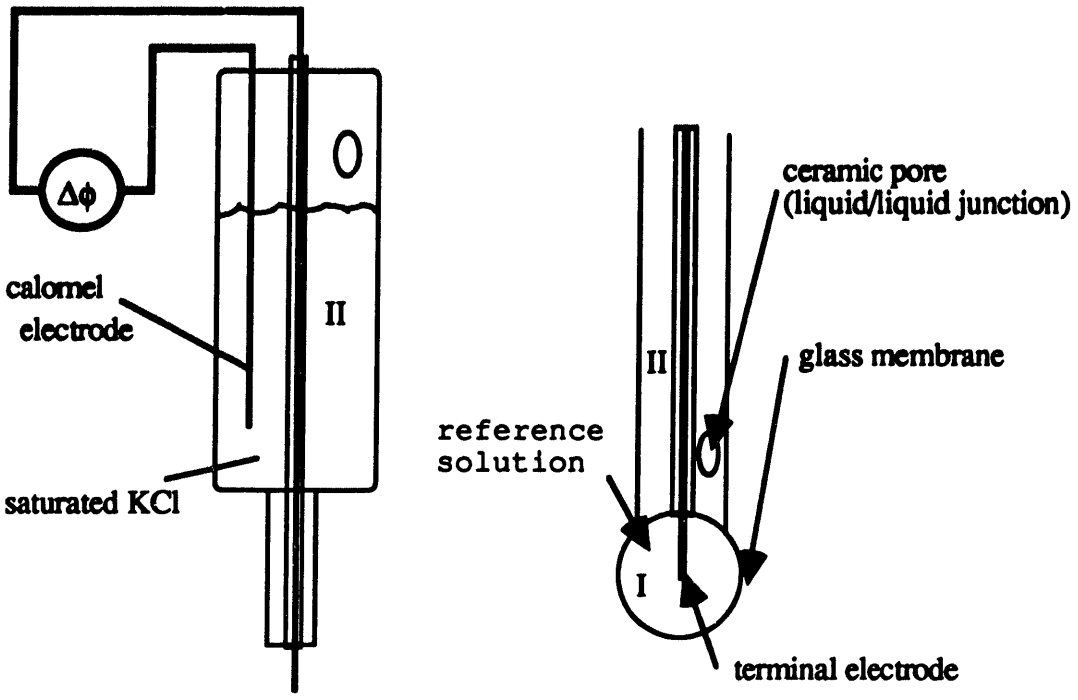


Figure 3.3 -- *pH* electrode. The upper portion of the electrode is on the left and the lower portion on the right. The two portions are connected. Compartment II is continuously connected, as is the central wire which terminates in the terminal electrode.

3.2.2 The Glass Membrane and $\Delta\phi_{glass}$

Ideally, the glass membrane functions as a semipermeable barrier which freely allows passage of H^+ ions between regions I and III, while at the same time blocking the passage of all other chemical species. In equilibrium the chemical potential of exchangeable particles be equal, i.e.,

$$\mu_I = \mu_{III} \tag{3.4}$$

Referring to chapter 2, this equality can be written as,

$$\mu^{o}_{I}(P,T)+kT\ln(f_{I}\chi_{I}) = \mu^{o}_{III}(P,T)+kT\ln(f_{III}\chi_{III})+e\Delta\phi_{glass} \quad (3.5)$$

where μ^{o} is the standard part of the chemical potential, f is the activity coefficient as defined in chapter 2, χ is the proton concentration, e is the fundamental charge, k is Boltzmann's constant, T is the temperature, and $\Delta\phi_{glass}$ is the potential difference across the glass membrane. Both of the standard chemical potentials in equation 3.5 are constants. Likewise, because the pH in the reference solution in compartment I is constant, the hydrogen ion activity coefficient and concentration are constant in this compartment. Noting that the definition of pH is,

$$pH = -\log_{10}(f\chi) \quad (3.6)$$

equation 3.5 can be written as,

$$\Delta\phi_{glass} = \frac{2.3\ kT}{e}\ pH_{III} + C \quad (3.7)$$

where C is a constant with contributions from the standard chemical potentials in both compartments I and III as well as from the pH in compartment I. This equation for the

potential difference at a junction is often called the Nernst equation[12].

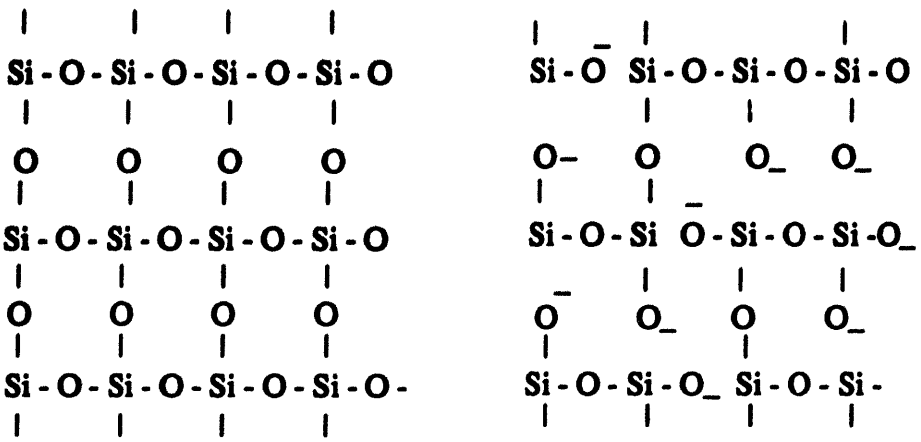


Figure 3.4 -- Silicon oxide in glass. The left picture represents an ideal lattice without impurities. The right picture represents a real lattice with numerous defects. These defects carry negative charge and facilitate the flux of cations through the glass membrane.

The physical basis for the selective permeability of the glass membrane can be understood in terms of the molecular structure of glass. Silica forms the major constituent of the glass membrane. In the absence of defects, four oxygens bind each silicon to form an interlinking network (see figure 3.4). The oxygens which bridge silicons are called binding oxygens. Non-binding oxygens are present, and carry a negative charge which facilitates their interaction with cation impurities. The effect of the non-binding oxygens is

[12]Bockris, J. O.; Reddy, A. K. N. *Modern Electrochemistry;* Plenum Press: New York, 1970.

to provide a conduit through the network for cations. Numerous other oxides can form glass as well, e.g., NaO_2 and CaO. In general, glass is a mixture of various oxides. The proportion of each oxide influences the permeability of the glass to different cations.

3.2.3 The Liquid Junction and $\Delta\phi_{liquid\ junct}$

There are several types of liquid junctions used in electrochemistry. The type used in this experiment is referred to as a "restrained flow" junction. A porous ceramic plug provides contact between regions II, saturated KCl, and III, the solution being studied. (See figure 3.3.) To prevent corruption of region II by back diffusion across the plug a constant small flux of KCl out of the pore was maintained by gravity. As a result of this flux compartment II periodically needed to be refilled with saturated KCl.

In order to explain the potential that arises across a liquid junction, several concepts and definitions must first be introduced.[13] The mobility of an ionic species, u_i, is a measure of its velocity in response to an applied electric field. It can be written as,

---

[13]Bard, A.; Faulkner, L. *Electrochemical Methods;* John Wiley & Sons: New York, 1980.

$$u_i = \frac{v}{E} = \frac{|z_i|e}{6\pi\eta r} \tag{3.8}$$

where v is the terminal velocity, η is the viscosity of the solution, e is the fundamental charge, E the applied field, z the valence, and r the ionic radius. The terminal velocity was determined by assuming frictional drag force as in the case of laminar flow past a sphere.

The conductivity, κ, of a solution can be expressed in terms of the mobility of the constituent species,

$$\kappa = F \sum |z_i| u_i C_i \tag{3.9}$$

where F is the Faraday constant, C_i is the concentration of the species, and i and other symbols are as previously defined.

The last definition needed is that of the transference number of an ionic species in solution. This is the fractional contribution of that species to the conductivity of the solution, i.e.,

$$t_i = \frac{|z_i| u_i C_i}{\sum |z_j| u_j C_j} \tag{3.10}$$

Imagining that the junction is divided into infinitesimal cells, the change in free energy associated with the flux of ions through a cell can be written in terms

of the transference number as,

$$dG = \sum_i \frac{t_i}{z_i} d\mu_i \qquad (3.11)$$

where $d\mu_i$ is the chemical potential change associated with species i crossing the cell. Integrating across the pore or junction gives a result that must be zero since the two regions are in equilibrium, i.e.,

$$\sum_i \int_{II}^{III} \frac{t_i}{z_i} d\mu_i = 0 \qquad (3.12)$$

The chemical potential of species i can be written as,

$$\mu_i = \mu^o_i + kT \ln\alpha_i + z_i e\phi \qquad (3.13)$$

where the activity of species is $\alpha_i = f_i\chi_i$, and other symbols are as previously defined. Using equation 3.13, the differential in equation 3.12 can be expressed to give[14],

$$\sum_i \int_{II}^{III} \frac{t_i}{z_i} kT\, d\ln\alpha_i + \left(\sum_i t_i\right) e \int_{II}^{III} d\phi = 0 \qquad (3.14)$$

or,

---

[14]Albert, A.; Serjeant, E. P. *The Determination of Ionization Constants;* Chapman and Hall: London, 1984.

$$\Delta\phi_{liquid\ junct.} = -\frac{kT}{e}\sum_i \int_{II}^{III} \frac{t_i}{z_i}\, d\ln\alpha_i \qquad (3.15)$$

where t_i, z_i, and α_i are the transference number, valence, and activity of species i of ion, respectively. An exact solution for the liquid junction potential is not possible since it would require the exact activity profile of the ions across the junction.

Several approaches have been devised to overcome this problem. Henderson[15,16] evaluated equation 3.15 by assuming that within the junction ionic concentrations are equivalent to activities, and that the concentration of each ion varies linearly between the boundaries of the junction. Equation 3.15 is then integrable and yields the following result for the liquid junction potential,

$$\Delta\phi_{liquidjunct} = \frac{kT}{e}\frac{\sum_i \frac{|z_i|u_i}{z_i}(C_i^{III}-C_i^{II})}{\sum_i |z_i|u_i(C_i^{III}-C_i^{II})}\ln\left(\frac{\sum_i |z_i|u_i C_i^{II}}{\sum_i |z_i|u_i C_i^{III}}\right) \qquad (3.16)$$

where C is concentration and other symbols are as defined previously. The sum is over all species of ions.

This equation provides justification for the use of K^+ as the counter ions in the saturated KCl solution within the

---

[15]Henderson, P. *Z. Phys. Chem.* **1908**, 59, 118.
[16]Eisenman, G. *Glass Electrodes for Hydrogen and Other Cations;* Marcel Dekker: New York, 1967.

electrode. Because K^+ and Cl^- have virtually identical mobilities, the use of K^+ as the counter ion resulted in a substantially lower junction potential than would be expected if another cation were used. More importantly, it is also evident from equation 3.16 that the dependence of the liquid junction potential on *pH* is extremely weak. To see this explicitly, note that the concentration of hydrogen ions is much less than the concentration of electrolyte. Because of this the logarithm is essentially constant. With the logarithm factor taken to be constant there is only one term which is dependent on the *pH* of the solution being studied. To show this term, equation 3.16 can be written as,

$$\Delta\phi_{liquidjunct} = \frac{kT}{e} \frac{u_{H^+}\,[H^+]^{III}}{\sum_I |z_i| u_i (C_i^{III} - C_i^{II})} \ln\left(\frac{\sum_I |z_i| u_i C_i^{II}}{\sum_I |z_i| u_i C_i^{III}}\right) + K \quad (3.17)$$

This term would be maximum when there are no charged species in region II other than protons, and when the *pH* in that region is 7. In that situation equation 3.17 becomes,

$$\Delta\phi_{liquidjunct} << \Delta\phi_{max} \approx 8.1 \frac{kT}{e}\; 10^{-pH_{III}} + K \quad (3.18)$$

When compared with equation 3.7, it is evident that the *pH* dependence of the liquid junction potential is orders of magnitude smaller than the *pH* dependence of the glass membrane potential. Accordingly, it will be considered to be

constant for a given experiment in which *pH* alone is varied.

In summarizing sections 3.2.1 to 3.2.3, it has been shown that the total potential difference measured by the glass membrane combination electrode used in these experiments can be written, neglecting the very weak *pH* dependence of the liquid junction potential, as,

$$\Delta\phi \approx \Delta\phi_{glass} + C \tag{3.19}$$

or,

$$\Delta\phi = \frac{2.3\ kT}{e}\ pH_{III} + C \tag{3.20}$$

Equation 3.20 states that the potential difference measured by the electrode is proportional to the *pH* of the solution being studied. To calculate the constant in equation 3.20 is difficult. It is however, not necessary to do so in order to measure *pH*. The constant is determined by calibration of the electrode with solutions of known pH.

In reality, there is a slight *pH* dependence of *C* with a long associated time constant. This results in a drift in the response of the electrode when moved to significantly different *pH* conditions. This drift is illustrated in figure 3.5 where the electrode was abruptly moved from *pH* 7 to *pH* 1.2. As figure 3.5 shows, the time constant associated with

equilibration is on the order of 1.5 hours. Several measures were used to overcome the problem of electrode drift. They are discussed in section 3.4

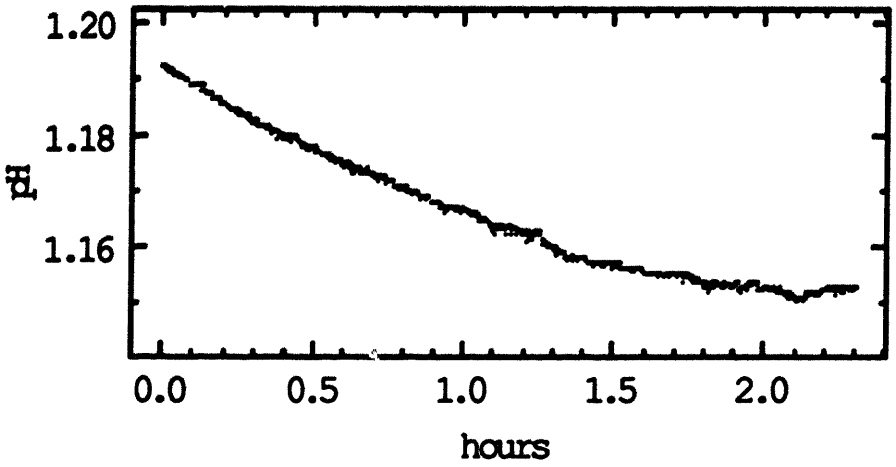


Figure 3.5 -- Electrode drift when the electrode was moved from *pH* 7 to approximately *pH* 1.2.

3.3 Handling and Preparation of the Protein Samples

The proteins studied were bovine lens γ_{II}, γ_{IIIa}, γ_{IIIb}, and γ_{IV}. Olutayo Ogun, a laboratory technician, purified these proteins from bovine lenses using a technique detailed in an in-house document entitled *The Gamma Factory*. The proteins were provided by Mr. Ogun at concentrations of about 1 mg/ml in buffered aqueous solution. The γ_{II} and γ_{III} solutions were provided in 25 mM ethanolamine buffered aqueous solution at

pH of 8.8. The γ_{IV} solution was provided in a mixture of 0.2M sodium acetate and 0.2M tris acetate buffered aqueous solution at *pH* of 6. All solutions initially contained 0.02% by weight sodium azide to inhibit bacterial growth. The hen egg lysozyme, which was used for control experiments, was purchased from Sigma Corporation in lyophilized form. It was then dissolved, filtered and thereafter treated in the same manner as the crystallins (see description below).

In order to successfully perform the experiments, the protein had to be dialyzed in order to remove buffer and impurities from the solution. In the standard dialyzing procedure, the protein is placed in a "diaflo" porous bag. Smaller molecules equilibrate with an aqueous reservoir and by repeatedly exchanging the reservoir their concentration is exponentially decreased. The antibacterial agent used as a preservative, sodium azide, was found to be a powerful buffer. Because of this, it too had to be removed from the solution. The standard procedure of allowing equilibration with a sequentially changed water reservoir could not be used because it requires several days. During this time bacterial growth could begin. To overcome this problem the protein solutions were sequentially concentrated and diluted. An Amicon pressurized concentrator was used with tank Nitrogen at 50 p.s.i.. The dilution factor of each stage of this procedure was approximately 10:1. This procedure was

repeated until at least a one million fold reduction in buffer and sodium azide concentration was achieved. Since the starting concentration of buffer was typically 0.2 M, and the starting concentration of sodium azide was .03 M, this provided a 100 fold safety margin (i.e., the concentration of the proteins was expected to be at least 100 times that of the buffers and 1000 times that of the azide at the time of the experiment).

After purifying the solution, the desired salts were added. "Blank" water solutions were prepared with identical salt conditions to those in the protein sample. The protein solutions, the acid or base to be used, and the "blank" water solutions were then degassed for two hours in a vacuum desiccator while being stirred with magnetically driven stir bars. After degassing, the vacuum desiccator was flooded with nitrogen.

Samples of the protein solution were removed at this point and their concentrations determined. It was necessary to know protein concentration in order to convert the raw data into the form of charge versus *pH* as described in chapter 4. The protein concentrations were determined by measuring the absorption of ultraviolet light at 280 nm. The extinction coefficients used for these concentration determinations are taken from the literature, and are

reproduced in table 3.1.[17,18] They were obtained by measuring the extinction and volume and subsequently determining the weight after lyophilizing the sample.

Table 3.1

γ_{II}	2.18
γ_{IIIa}	2.33
γ_{IIIb}	2.11
γ_{IV}	2.25
lysozyme	2.64

The acids and bases used as titrant were mixed from standard "acculyte" concentrates. Careful dilutions were carried out with volumetric flasks whose accuracy had been verified by weight of water measurements. Because of the possibility of carbon dioxide contamination, the base was mixed just prior to experiments, the base container was flooded with nitrogen after preparation, and the concentration of base was verified by titration against standard acid.

The volume of protein solution was determined by pipette, the accuracy of which had been determined to be within 0.5% by consecutive weight of water measurements. In typical experiments, between 3 and 10 ml of protein solution

---

[17]Broide, M. L.; et. al. *Proc. Natl. Acad.* **1991**, 88, 5660
[18]Taratuta, V.; et. al. *J. Phys. Chem.* **1990**, 5, 2140

at concentrations of approximately 10^{-4} M was used. Both the protein-containing flask and the acid or base reservoir were placed in their respective water baths while being kept under flowing nitrogen throughout the experiment. The nitrogen was first bubbled through distilled water and 5 N KOH to insure that it was saturated with water and depleted of carbon dioxide. Fifteen minutes was allowed for temperature equilibration before beginning an experiment. During this time the *pH* electrode was calibrated. The electrode was calibrated at either *pH*s 4 and 7 or 7 and 10 depending on whether the experiment was to be conducted in the acidic or basic range, respectively. Prior to beginning the experiment the *pH* of the protein solution was adjusted to the desired starting point by the addition by hand of small aliquots of .1 N acid or base. The volumes added, though small, were recorded and used as corrections to the starting volume.

In the case of experiments outside of the middle *pH* range (4 to 7), additional measurements were needed to allow for corrections to account for the non-ideal behavior of the electrode. These measurements consisted of titration runs with water and electrolyte alone. These "blank" water titration experiments were collected both before and after the titration experiment with protein to allow for characterization of the electrode drift as well as the electrode response. The manner in which these water titration runs were used to correct the data is discussed in

chapter 4. The possibility of protein aggregation or precipitation was excluded by repeat spectrophotometric concentration measurements at the end of experiments.

3.4 Experimental Uncertainties

Several uncertainties persisted in this procedure: (1) uncertainty in the extinction coefficients of the crystallins, and corresponding uncertainties in the concentration of protein (see table 3.1); (2) uncertainty in volumes, which were determined by consecutive weight of water measurements to be on the order of a fraction of a percent; (3) uncertainties in electrode calibration and non-ideal electrode behavior which was corrected for by the use of water titration runs (see chapter 4); (4) uncertainties resulting from carbon dioxide uptake in both the protein solutions and the base titrant (Every effort was made to avoid this complications by degassing and performing the experiments under nitrogen. Nonetheless, the possibility of some contamination with carbon dioxide, particularly in the extreme basic range, cannot be entirely excluded.); (5) uncertainties in acid or base titrant concentration (These uncertainties are given by the manufacturer (Acculyte) as < 0.1%. Some additional slight uncertainty may accrue by virtue of the dilutions.); and (6) electrode drift (See the

discussion in section 3.2.2 on the physical origin of electrode drift and the associated time constant). The significance of electrode drift in effecting the data, and the measures taken to minimize electrode drift, warrant special consideration.

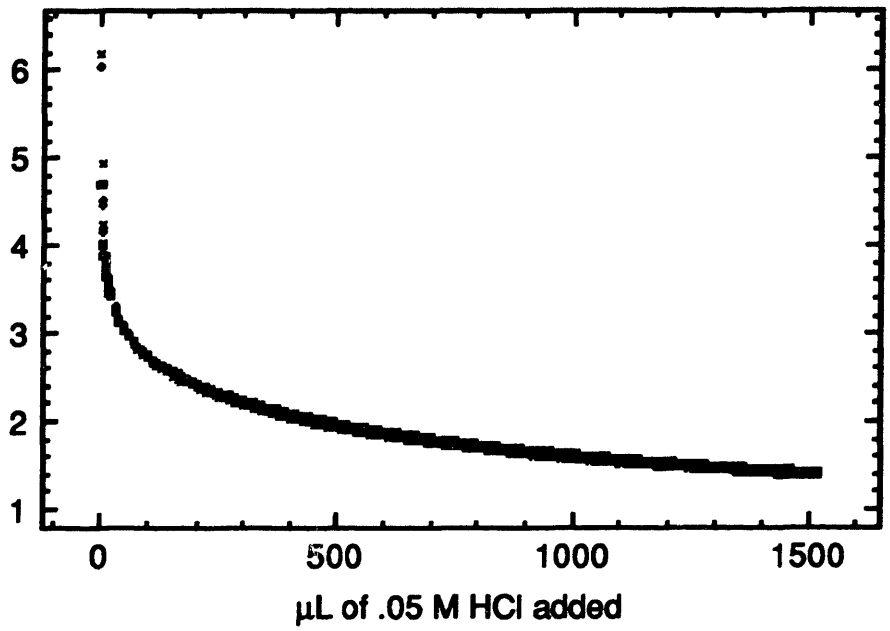


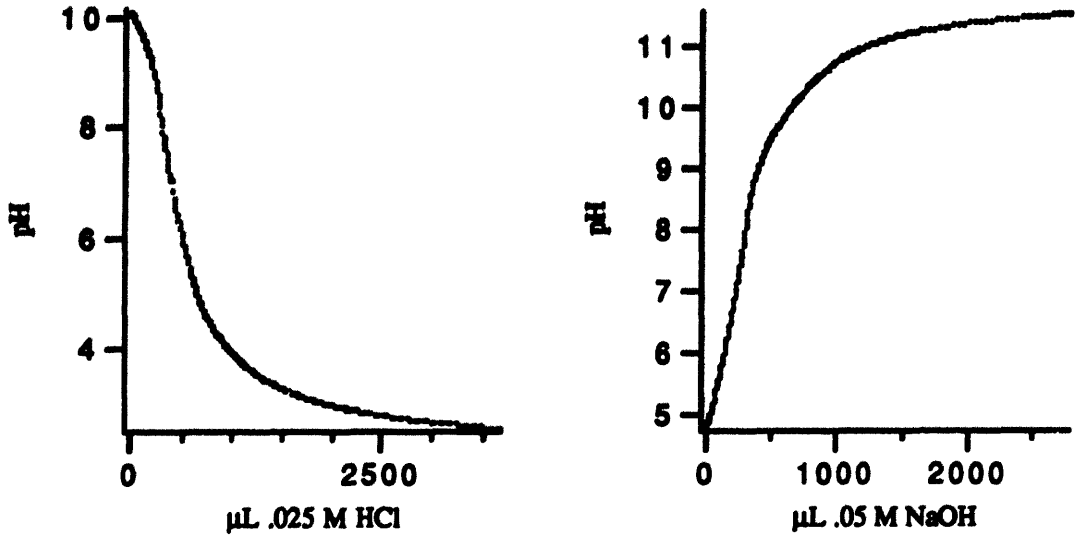
Figure 3.6 -- Four consecutive titration experiments with identical water samples. The electrode had been preconditioned as described in the text.

As discussed in section 3.2, the effect of electrode drift was enhanced by rapid and large changes in the *pH* environment. In an attempt to minimize electrode drift, several measures were undertaken. Most importantly, the electrode was conditioned to the *pH* environment in which it

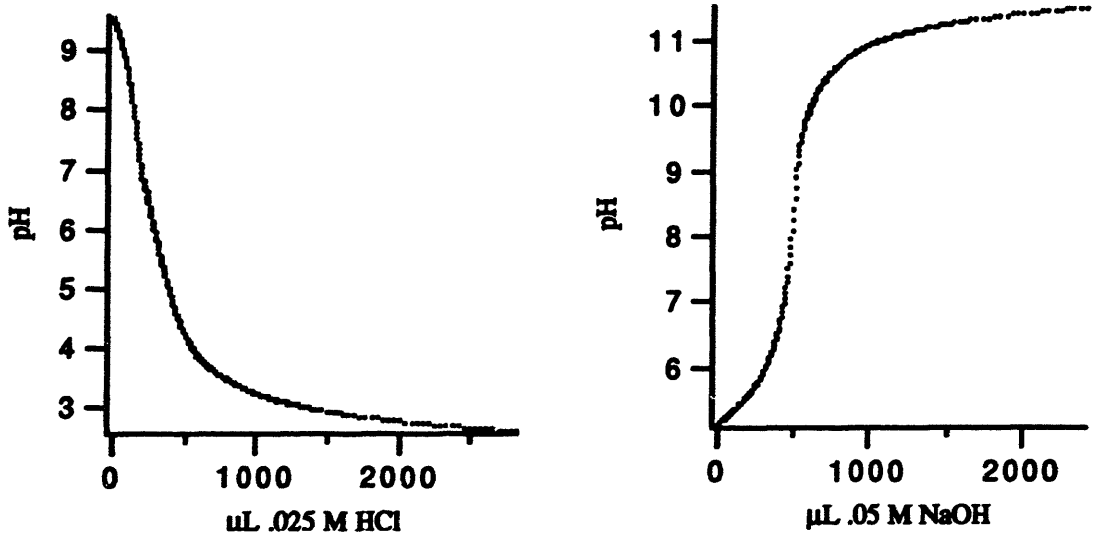
would be used by presoaking. Also, the range of *pH* to be explored in a single experiment was restricted. The success of these measures is evident in figure 3.6, where 4 consecutive sets of raw water titration data are presented. The samples are identical and the electrode was soaked in solution at *pH* 2.5 for 2 hours prior to collecting the data. The effect of electrode drift was extremely small, resulting in no detectable change in these consecutive titration experiments with identical samples.

3.5 Raw Titration Data

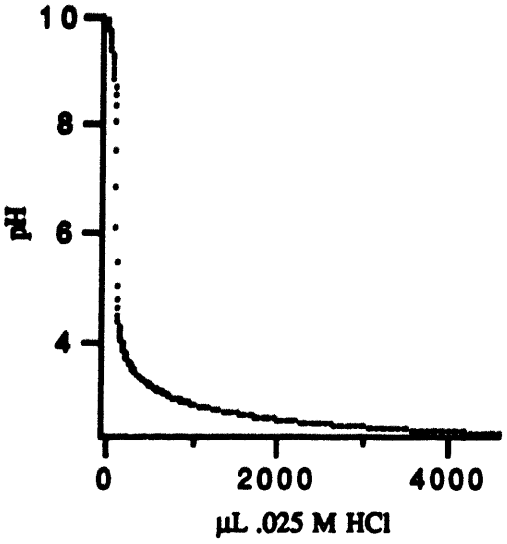
A representative sample of the raw titration data is reproduced below in figures 3.7 to 3.12. In this form the data is of little use. The appearance of a particular curve depends on the starting volume of protein, on the concentration of protein, and on the concentration of base or acid, as well as on the intrinsic proton binding characteristics of the protein being studied. Converting this raw data into a form representing the intrinsic proton binding properties of the protein, i.e. charge versus *pH*, is the subject of the next chapter.

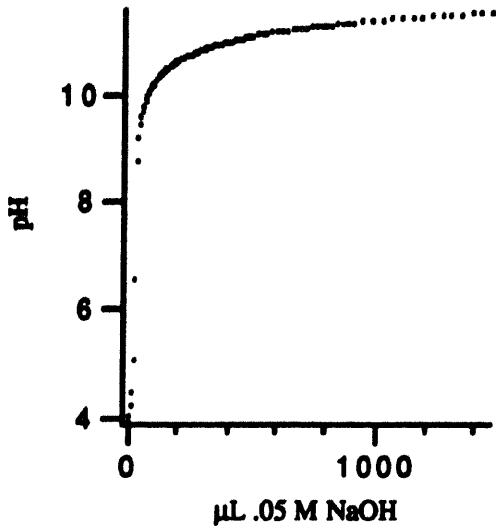


Figures 3.7 & 3.8 -- 10 ml γ_{IIIb} in 0.1M KCl raw titration data in both acidic and basic range experiments.



Figures 3.9 & 3.10 -- Raw titration curves of 10 ml γ_{IV} in .01 M KCl.





Figures 3.11 & 3.12 -- Raw "blank" titration curves of water with 0.1 M KCl.

Chapter 4 - Determining Protein Charge as it Depends on *pH*

4.0 Introductory Remarks

The raw acid-base titration data from chapter 3 depends on the details of a particular experiment, such as volume of solution, concentration of acid or base, and concentration of protein. In this form the data offers little insight into the proton binding properties of a protein. Ideally, the data should be in a form which is independent of these parameters, i.e. net charge versus *pH*. This chapter explains two methods of transforming the data into charge as a function of *pH*.

The first of these methods involves keeping track of the number of protons or hydroxyls added to the protein solution as titrant. By correcting for the binding of hydroxyls and protons to form water, the number of protons expected to be free in solution can be determined in terms of the acid or base added. This is compared to the number of protons determined to be in solution by *pH* measurement. The

difference between the measured and expected number of protons in solution is taken to be the number of binding events between protons and proteins. This method assumes ideal response of the electrode because it is on the basis of *pH* measurement that the number of protons in solution is determined. This method provides good results over a restricted *pH* range (3 to 10).

To extend the *pH* range over which the data can be transformed to charge versus *pH*, it is necessary to use blank titration experiments in which water and electrolyte are titrated without protein. The blank titration experiments had the same salt conditions (ionic strength and salt identity) as their corresponding protein solution titration experiments. In addition, the volume of the blank solutions was the same as that of their protein solution counterparts. Neither blank water nor protein contained any buffer, since this would interfere with the determination of proton binding. The number of protons or hydroxyls is the difference in the amount of acid or base added to attain the same *pH* in the protein solution and the blank solution is the number of proton binding events as a function of *pH*. This method is much more powerful than the one described above since the electrode response is not used to determine the number of protons in solution.

Both of these methods are presented in detail in this chapter. After doing so, the resulting charge versus *pH*

curves for the γ-crystallins are presented. The effects of electrolyte concentration and electrolyte identity on the proton binding curves are also examined. The propagation of experimental uncertainties is discussed in detail in this chapter.

4.1 - Charge versus *pH* Transformation Assuming Ideal Electrode

This method of transforming the raw titration data from chapter 3 assumes an ideal electrode, with the measured *pH* being equal to $-log_{10}([H^+])$ over all *pH* ranges. In principle, in order to express the raw data as charge versus *pH*, it is only necessary to keep track of the number of protons (or hydroxyls) added to the solution and the *pH*. It can be assumed that protons which do not register as a *pH* change are bound to protein. In order to determine the average number of protons binding per protein, the total number of binding protons is divided by the number of protein molecules in solution. The equilibrium between protons, hydroxyls, and water molecules must also be taken into account, i.e.,

$$H^+ + OH^- \leftrightarrow H_2O \qquad (4.1)$$

Each H_2O can be thought of as a proton plus a hydroxyl. Therefore, protons are present either as water molecules, free in solution, or bound to protein.

Protons were added during the course of an experiment either in the form of water molecules in the titrant or as free hydrogen ions in the titrant. Let H be the number of free protons in solution, OH the number of free hydroxyls in solution, H_2O the number of water molecules in solution, c_H the concentration of acid in titrant, c_{OH} the concentration of base in titrant, ΔV the volume of titrant added, ΔH_2O the number of water molecules added in the titrant, P the number of protein molecules in solution, and ν the average number of protons bound to each protein. The subscript o means the initial values of a particular parameter, and brackets ($[\ ]$) signify concentration as opposed to number. Conservation of hydrogens implies,

$$(H_2O_o + H_o + \nu_o P) + \Delta H_2O + c_H \Delta V = H_2O + H + \nu P \qquad (4.2)$$

In principle, this is enough information to determine $\nu - \nu_o$ as a function of pH. Most hydrogen is present as H_2O. Since this hydrogen is extraordinarily more abundant than the protons bound to protein or free in solution, the slightest fractional error in the number of water molecules would

invalidate the result. The resolution to this problem can be obtained from the conservation of hydroxyls.

The same reasoning used for equation 4.2 gives rise to an equation for the conservation of hydroxyls,

$$H_2O_o + OH_o + \Delta H_2O + c_{OH}\Delta V = H_2O + OH \qquad (4.3)$$

where it is assumed that there is no binding of hydroxyls to protein. This second equation allows for the elimination of terms involving the number of water molecules, i.e., subtracting equation 4.2 from equation 4.3 gives,

$$(V - V_o)\ P = (c_H - c_{OH})\ \Delta V + H_o - H + OH - OH_o \qquad (4.4)$$

Neglecting non-unity of activity coefficients (see section 2.2), the ionic product of water can be stated as,

$$[H^+]\,[OH^-] = 10^{-14} \qquad (4.5)$$

where the brackets indicate molar concentration. This allows expression of equation 4.4 as,

$$(V-V_o)=\frac{1}{V_o[P]_o}\left((c_H-c_{OH})\Delta V+V_o(10^{-pHo}-10^{pHo-14}) -V(10^{-pH}-10^{pH-14}) \right) \qquad (4.6)$$

The ionic product of water holds for the titrant as well. When titrating with acid the concentration of base is sufficiently small so that one can neglect c_{OH} in the titrant. Likewise, when titrating base one can neglect c_H. Equation 4.6 can be further simplified depending on the *pH* range of a particular experiment and whether acid or base is added. For acidic range titration the appropriate transformation becomes,

$$(V-V_o) \approx \frac{1}{[P]_o}\left(c_H \frac{\Delta V}{V_o} + 10^{-pHo} - \frac{V}{V_o} 10^{-pH}\right) \qquad (4.7)$$

Alternatively, for basic range titration it becomes,

$$(V-V_o) \approx \frac{1}{[P]_o}\left(-c_{OH} \frac{\Delta V}{V_o} + \frac{V10^{pH-14}}{V_o} - 10^{pHo-14}\right) \qquad (4.8)$$

This procedure, described by equations 4.7 and 4.8, produces charge versus *pH* curves that have erroneous asymptotes at both *pH* extremes. This can be seen in figure 4.1 where the charge versus *pH* curves of glycine and acetic acid have been determined using this method. Both of these substances should exhibit only one binding event. The curves should exhibit flat regions as one approaches the *pH* extremes since no binding is expected there. Instead, there is an exponentially increasing error as the *pH* extremes are approached.

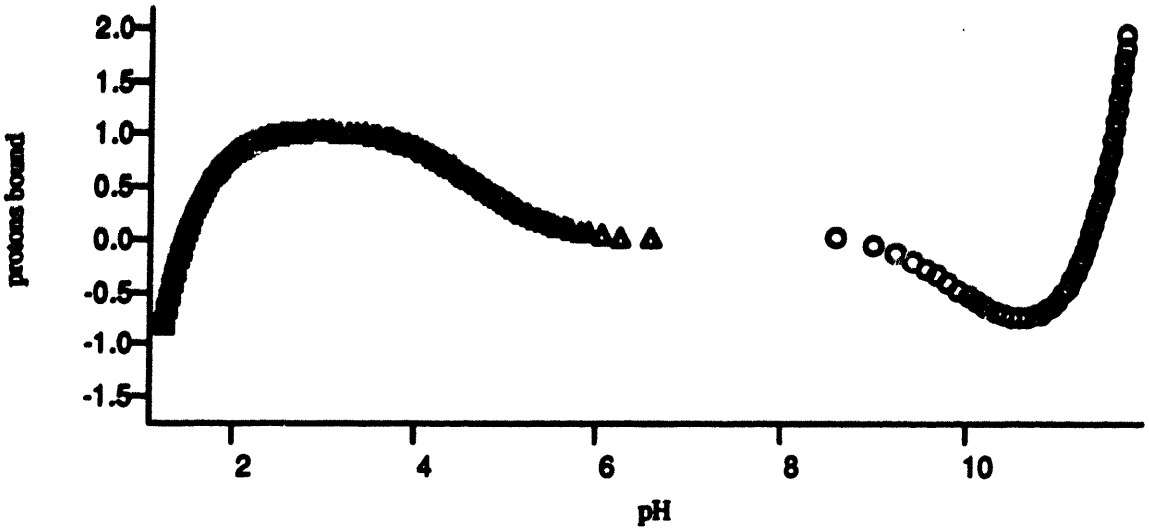


Figure 4.1 -- This figure demonstrates errors which occur in the *pH* extremes using the transformation described in section 4.1 which fails to correct for non-ideal electrode behavior and non-unity of the activity coefficients. The circles are titration data from glycine and the triangles are titration data from acetate.

The anomalous asymptotes seen in figure 4.1 are due to small errors in the electrode's calibration and deviation from unity of the activity coefficient of the proton. Using the transformation described by equation 4.6, however, these small errors in *pH* translate into large errors in the protein charge at more extreme *pH*s. This results from the logarithmic relationship between *pH* and concentration and can be appreciated by noting that at *pH* 2, a one percent error in the *pH* corresponds to an error of 10^{-4} M in proton concentration. Typical values for concentration of the proteins in solution are on the order of 10^{-5} M. In other words, a one percent uncertainty in *pH* at *pH* 2 represents an

uncertainty on the order of ten binding events per protein molecule.

To explicitly see the uncertainty of protein charge in *pH* extremes, equations 4.7 and 4.8 can be differentiated to show how *pH* errors propagate. In the acidic range,

$$\delta\nu = \frac{\partial\nu}{\partial pH}\,\delta pH = -\frac{\ln 10\ V\ 10^{-pH}}{V_o[P]_o}\,\delta pH \qquad (4.9)$$

Likewise, in the basic range,

$$\delta\nu = \frac{\partial\nu}{\partial pH}\,\delta pH = \frac{\ln 10\ V\ 10^{pH-14}}{V_o[P]_o}\,\delta pH \qquad (4.10)$$

In these equations $\delta\nu$ is the uncertainty in protein charge, δpH the uncertainty in *pH*, and other symbols are as previously described. Equations 4.9 and 4.10 demonstrate that the uncertainty in binding events increases exponentially for a given uncertainty in *pH* as one approaches the *pH* extremes. This problem is further compounded by increasingly non-ideal behavior of the electrode and non-unity of activity coefficients as the *pH* extremes are approached. The *pH* at which these errors become problematic depends on the concentration of the protein in solution. At the protein concentrations used in the experiments described in chapter 3 significant errors arose below *pH* 3 and above *pH* 10.

The transformation presented in this section illustrates how, in principle, charge versus *pH* can be derived from acid-base titration data alone. In practice, however, this approach is inadequate in the *pH* extremes. An improved transformation using titration curves of water to correct for the anomalous behavior of the electrode and for non-unity of the hydrogen ion activity coefficient is presented below.

4.2 Charge versus *pH* Transformation Correcting for Non-Ideal Electrode Behavior and Non-Unity of Activity Coefficients: the Use of Water Calibration Experiments

The most direct method with which to accurately determine charge versus *pH* involves the use of reference titration curves of water without protein. These "blank" water titration curves must be collected with identical volumes, salt conditions, titrant. The protein data file will be denoted by $pH_p(\Delta V)$, and the water file by $pH_w(\Delta V)$. These files are sets of data pairs corresponding to *pH* and the volume of acid or base which was added to attain the corresponding *pH*. Examples of such files for γ_{IV} in 0.1 M KCl are presented in figure 4.2.

It is first necessary to invert these files, writing them as $\Delta V_p(pH)$ and $\Delta V_w(pH)$, where the volume of acid or base added can now be plotted as a function of *pH* for both the

protein solution and the water. The amount of acid or base titrant required to achieve a certain *pH* in water is subtracted from the amount required to achieve the same *pH* in protein solution. To illustrate this step, the data files in figure 4.2 have been inverted and their difference taken in figure 4.3. Since both $pH_p(\Delta V)$ and $pH_w(\Delta V)$ are raw data files, it is necessary to interpolate between the data points in $\Delta V_w(pH)$ in order to construct this difference function. The difference in the volume of titrant times the concentration of acid or base represents the number of binding events to protein.

When the number of binding events is divided by the number of protein molecules in the solution the result is the number of protons binding a single protein molecule versus *pH*. These steps can be summarized in quantitative form for titration with acid as,

$$\nu - \nu_o = \frac{C_H}{V_o[P]_o}\left(\Delta V_p(pH) - \Delta V_w(pH)\right) \qquad (4.11)$$

and for titration with base as,

$$\nu - \nu_o = \frac{-C_{OH}}{V_o[P]_o}\left(\Delta V_p(pH) - \Delta V_w(pH)\right) \qquad (4.12)$$

In equations 4.11 and 4.12, ν is the number of protons bound per protein, ν_o is the initial number of protons bound per

protein, C_H and C_{OH} are the concentrations of acid and base titrant, V_o is the initial volume of the solution, and $[P]_o$ is the initial protein concentration.

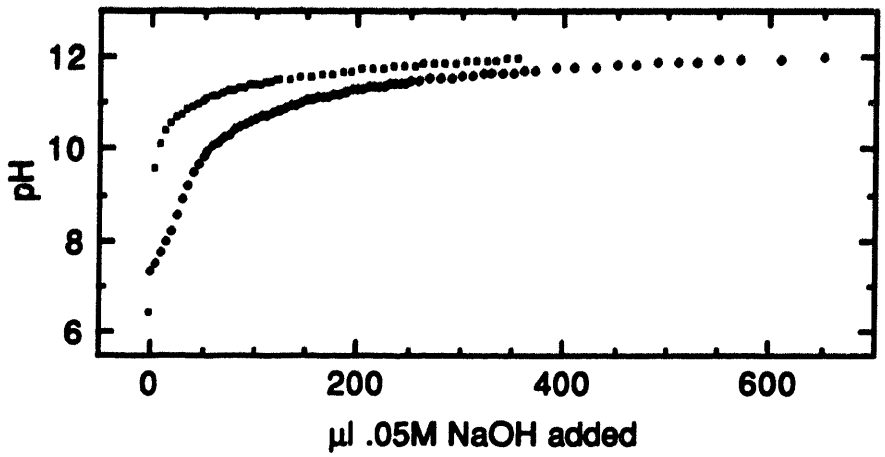


Figure 4.2 -- Raw basic range titration data from γ_{IV} (circles). Corresponding blank water titration data is presented in the same figure (dots).

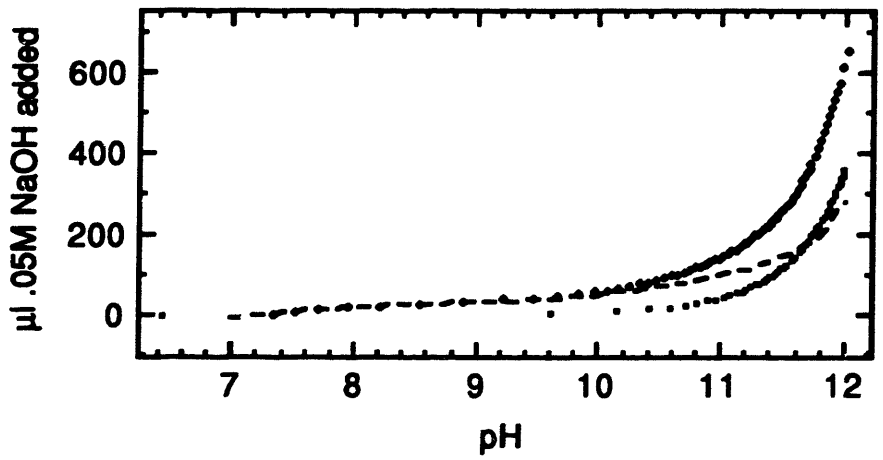


Figure 4.3 -- Raw titration data from figure 4.2 replotted with axes reversed. The difference between the protein titration data and the "blank" water data is also plotted with a dotted line.

Equation 4.12 is applied to the difference curve in figure 4.3 to reveal the charge versus *pH* curve for γ_{IV} in 0.1 M KCl in figure 4.4.

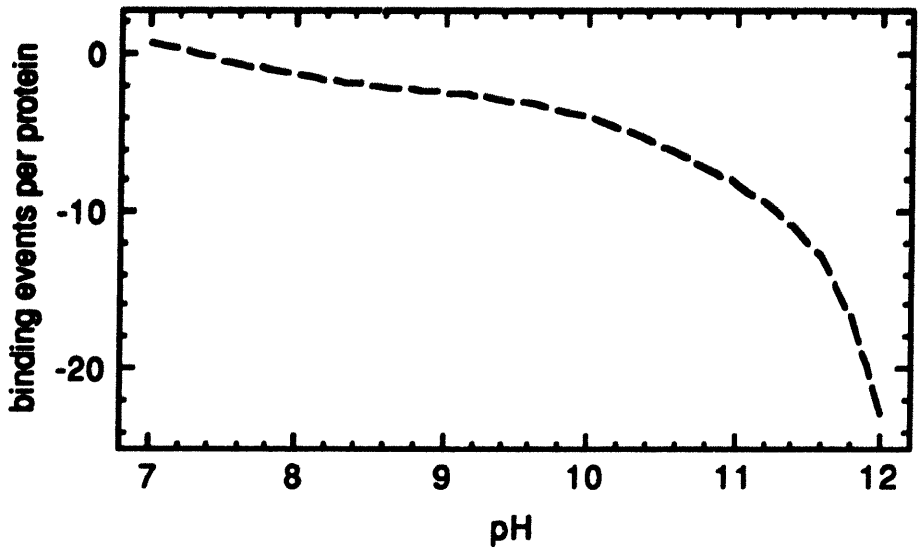


Figure 4.4 -- Basic range charge vs. *pH* curve for γ_{IV} in 0.1 M KCl. This curve was determined from figure 4.3 as described in the text.

4.2.1 Propagation of Uncertainties

The complicating factors of non-ideal electrode behavior, imperfect electrode calibration, and non-unity of the proton activity coefficient are all equally present in both the protein titration experiment and the blank water experiment. As a result, the calculated number of proton binding events is independent of these uncertainties.

One category of uncertainty not accounted for in this transformation is electrode drift. If the electrode were to drift significantly between the water "blank" experiment and the protein experiment, errors would be introduced into the difference between these two curves. This problem was addressed by repeating the water experiments, with one done immediately before the protein experiment and the other immediately after. In this fashion, the effect of electrode drift was demonstrated to give rise to negligible uncertainty. (For a more detailed discussion of drift refer to section 3.4.)

Additionally, attention must be paid to the propagation of experimental uncertainties in protein concentration, acid or base concentration, and solution volume. These uncertainties are propagated by differentiating equations 4.11 and 4.12 with respect each parameter and multiplying by its uncertainty. In the instance of protein concentration, differentiating 4.11 or 4.12 it is evident that,

$$\frac{\delta\nu}{(\nu-\nu_o)} \approx \frac{1}{(\nu-\nu_o)} \frac{\partial\nu}{\partial[P]_o} \delta[P]_o \approx \frac{\delta[P]_o}{[P]_o} \qquad (4.13)$$

i.e., a fractional error in protein concentration translates directly into a fractional error in the number of binding events. This is also the case for volume and acid or base concentration, i.e.,

$$\frac{\delta v}{(v-v_o)} \approx \frac{\delta V_o}{V_o} , \frac{\delta C_H}{C_H} , \frac{\delta C_{OH}}{C_{OH}} , \frac{\delta [P]_o}{[P]_o} . \qquad (4.14)$$

The total uncertainty in charge versus *pH* resulting from these experimental uncertainties can be written for acidic titration as,

$$\frac{\delta v}{(v-v_o)} \approx \sqrt{\frac{\delta V_o^2}{V_o^2} + \frac{\delta C_H^2}{C_H^2} + \frac{\delta [P]_o^2}{[P]_o^2}} \qquad (4.15a)$$

and for basic titration as,

$$\frac{\delta v}{(v-v_o)} \approx \sqrt{\frac{\delta V_o^2}{V_o^2} + \frac{\delta C_{OH}^2}{C_{OH}^2} + \frac{\delta [P]_o^2}{[P]_o^2}} \qquad (4.15b)$$

The fractional uncertainty of each of these parameters was approximately 1%, as described in chapter 3. This resulted in a fractional uncertainty in charge of about $\sqrt{3}$ x 1%.

4.2.2 Experimental Control: Reproducing Glycine, Acetic Acid, and Lysozyme Titration Data

To verify that this procedure was implemented properly it was tested on several systems with known acid-base titration curves. The first two were acetic acid and

glycine, and the third was the protein hen egg lysozyme. In all instances, the approach agreed with established results.

4.2.2.1 Glycine and Acetic Acid

The transformation method described in section 4.1 failed to produce accurate charge versus *pH* curves in the *pH* extremes. This was demonstrated in figure 4.1 for the cases of glycine and acetic acid. The method presented in section 4.2 was applied to acetic acid and glycine in figure 4.5. The predicted single binding events are manifest in each case, with no evidence of anomalous behavior to *pH*'s below 2 and above 12.

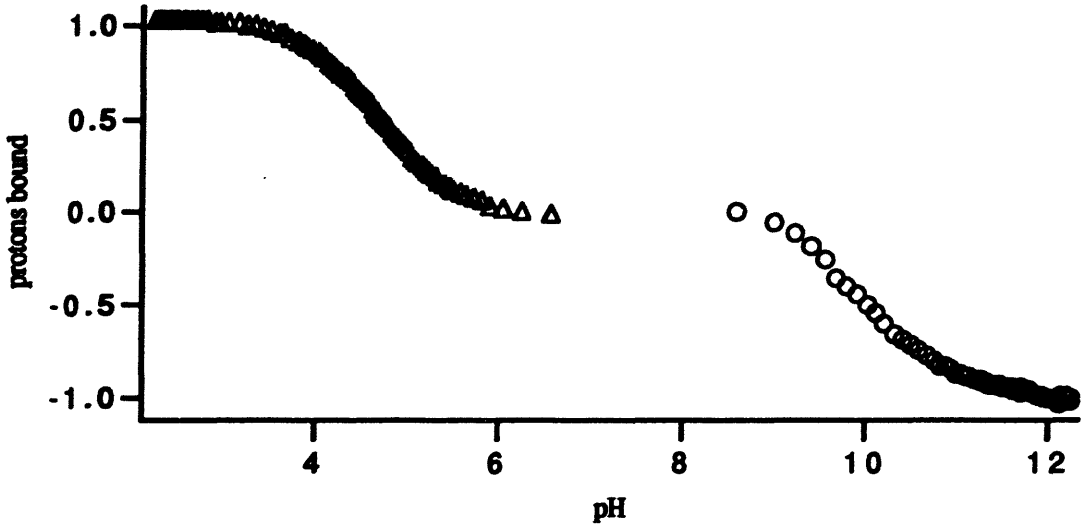


Figure 4.5 -- Experimental titration curves of acetate (triangles) and glycine (circles) analyzed using the method described in section 4.2.

4.2.2.2 Lysozyme

The next step was to compare results with prior work done on the titration of protein. One of the most extensively studied proteins is hen egg lysozyme. This final check verified the accuracy of the device and procedure used for studying the γ crystallins.

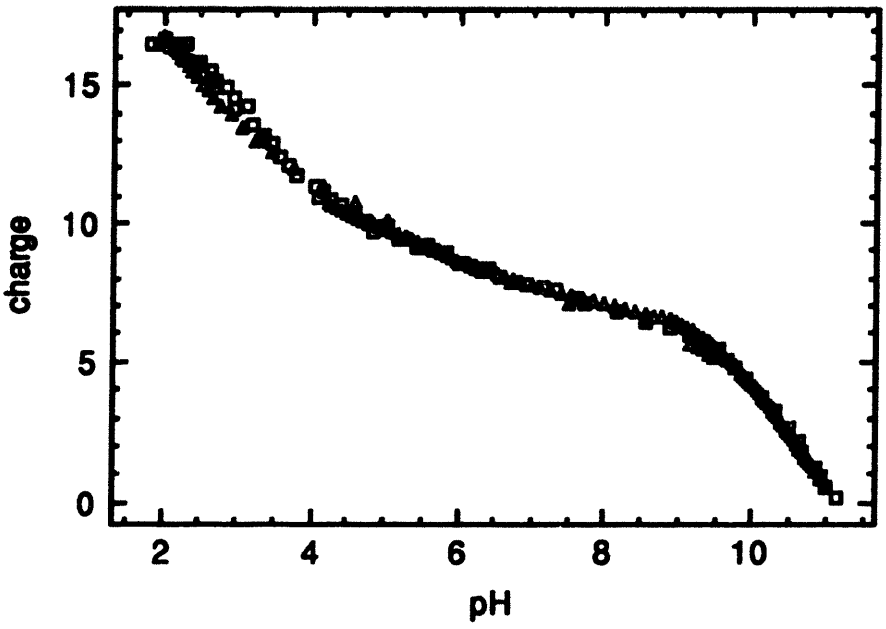


Figure 4.6 -- This is a comparison of titration data for lysozyme obtained in this laboratory (triangles) and by Kuramitsu[19] (squares).

[19]Kuramitsu, S.; Hamaguchi, K. *J. Biochem.* **1980**, 87, 1215.

The lysozyme titration experiment was performed in 0.1M KCl and the results are presented in figure 4.6 where they are compared to an earlier study by Kuramitsu.[20] As figure 4.6 demonstrates, the results obtained here were in excellent agreement with this prior study.

4.3 Isoelectric Points of the Crystallins -- Determining Offset

The transformations described in section 4.2 determine the change in protons bound to protein. This procedure in no way fixes the value of the charge. The offset can be chosen either so that the y-axis represents the total number of protons bound to the protein or so that it represents the net charge of the proteins. This later convention was used here. In order to implement this procedure, the net charge of the protein must be known at one *pH*. The isoelectric point, i.e. the *pH* at which the net protein charge is zero, is most easily determined. The isoelectric points had been previously determined for the γ-crystallins using gel electrophoresis.[21] These values are taken from the literature and reproduced in table 4.1.

---

[20]Kuramitsu, S.; Hamaguchi, K. *J. Biochem.* **1980**, 87, 1215.
[21]McDermott, M. J.; et. al. *Arch. Biochem. Biophys.* **1988**, 262, 609.

Table 4.1

Protein	Isoelectric Point
γ_{II}	7.8
γ_{IIIa}	7.8
γ_{IIIb}	7.4
γ_{IV}	7.8

4.4 γ-Crystallin Charge versus *pH* Data

Titration experiments were conducted for the bovine lens γ-crystallin fractions, γ_{II}, γ_{IIIa}, γ_{IIIb}, and γ_{IV} as described in chapter 3. The results were analyzed using the transformations described in section 4.2. Experiments were conducted in a variety of solution conditions to investigate the influence of electrolyte on proton binding.

4.4.1 Compilation of Data in 0.1 M KCl

Figures 4.7 through 4.11 present the charge versus *pH* data for the γ-crystallins in 0.1 M KCl. The offset of the y-axis was chosen based on the isoelectric points given in table 4.1 so that the y-axis represents net protein charge. Complete study of the proteins in 0.1 M KCl was of particular

interest, since this is similar to physiologic intracellular ionic conditions.

Because of the scale, it is difficult to read the exact charge from the figures 4.7 - 4.11. For future work, however, it may be necessary to know the charge to very high precision. For this reason, enlarged reproductions of figures 4.7 - 4.11 are presented in appendix G.

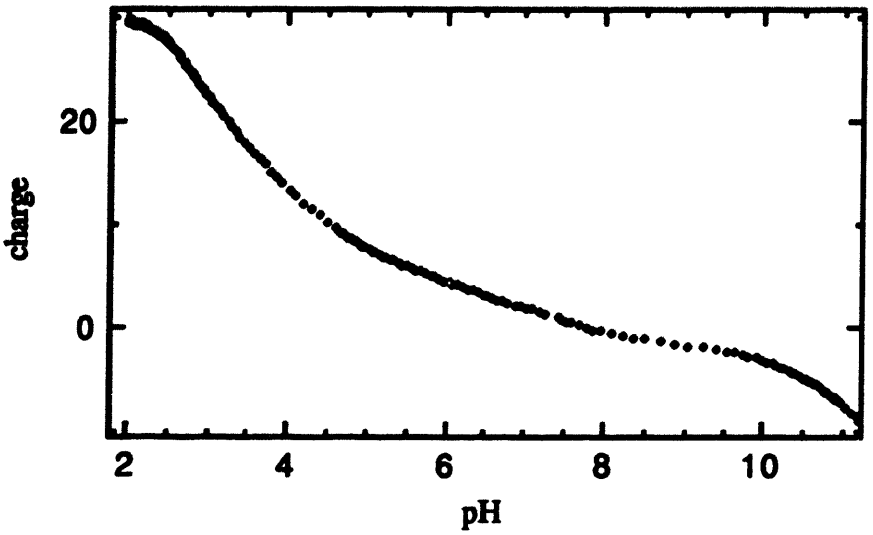


Figure 4.7 -- Charge versus *pH* titration curve of γ_{II} in 0.1M KCl. This is a composite curve from two different experiments, one with acidic titration and one with basic titration

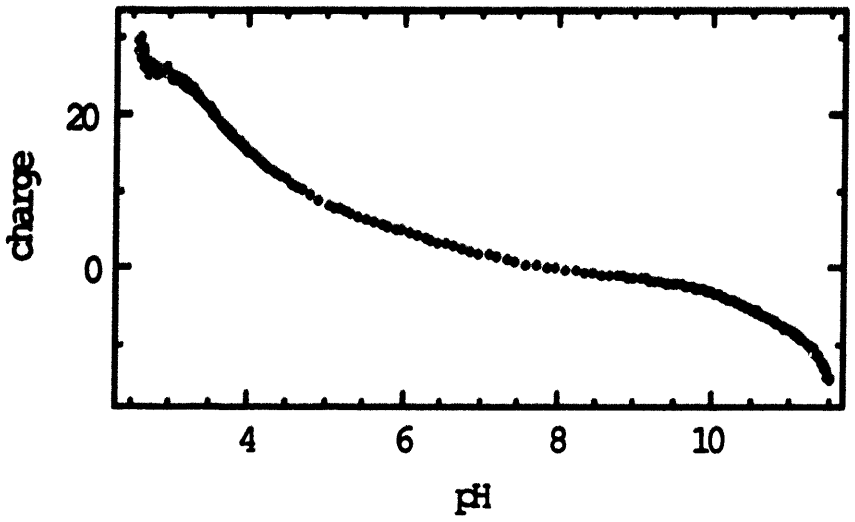


Figure 4.8 -- The charge versus *pH* titration curve of γ_{IIIa} in 0.1 M KCl. This is a composite curve from three different experiments.

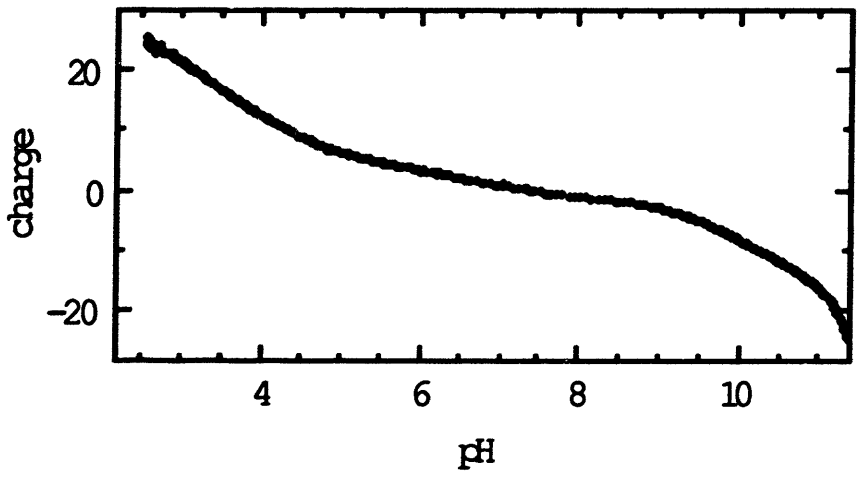


Figure 4.9 -- Charge versus *pH* titration curve of γ_{IIIb} in 0.1 M KCl

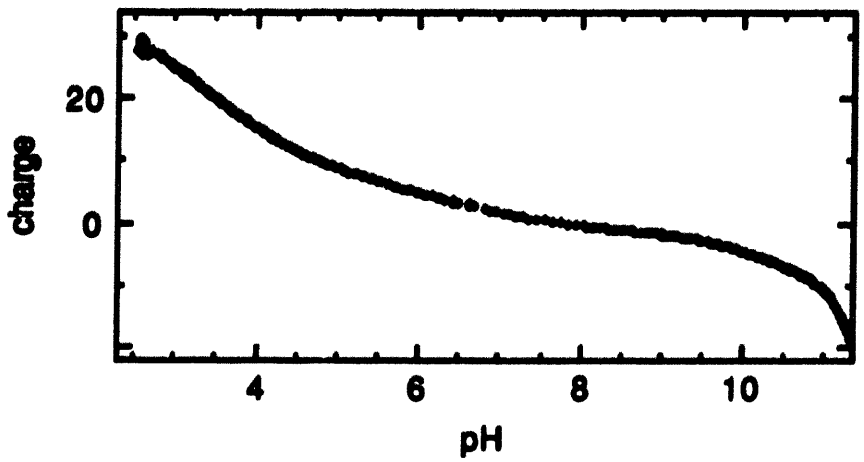


Figure 4.10 -- Charge versus *pH* titration curve of γ_{IV} in 0.1 M KCl. This curve is a composite of two experiments, one in the basic range and one in the acidic range. Note that since the collection of this data the fractionation of γ_{IV} into its sub-components, γ_{IVa} and γ_{IVb} has become practical. This curve was obtained before these subfractions were readily available.

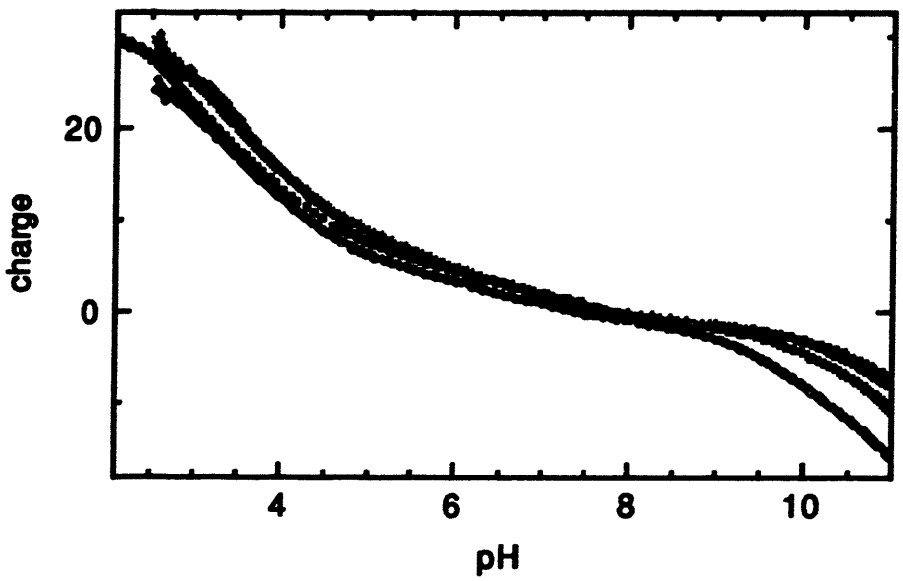


Figure 4.11 -- Collection of the last four figures into one. Here γ_{II} (square), γ_{IIIa} (plus), γ_{IIIb} (cross), and γ_{IV} (diamond) are plotted on one graph to illustrate the similarities and differences in acid-base titration behavior.

Qualitatively, the titration curves of the γ-crystallins are similar. With the exception of the basic range titration of γ_{IIIb}, there is little difference between the titration curves of these proteins. The basic range titration curve of γ_{IIIb} exhibits a steeper yet nearly constant slope in the *pH* range between nine and eleven.

In general, the titration curves of these proteins are typical when compared to those of other proteins. There is the usual flattening of the middle *pH* range when compared to the denatured state. This is universally observed in globular proteins, and represents the effects of electrostatic interactions on the *pK*s of the titrateable residues. This phenomenon will be discussed in chapter 5.

4.4.2 Effect of Salt Identity on Titration Curves of the γ Crystallins

The effects on the titration curves of γ_{IIIb} and γ_{IV} which resulted from varying the electrolyte identity were studied. Figure 4.12 and 4.13 reproduce the results of several experiments in which the identities of both cation and anion were changed. All of these experiments were conducted with electrolyte at 0.1 M. The salts investigated were potassium chloride, sodium chloride, and sodium bromide in the acidic

range for γ_{IIIb}, and potassium chloride and sodium chloride in the basic range for γ_{IV}. No appreciable change in the titration characteristics of either protein resulted from these changes in solvent.

4.4.3 Effect of Ionic Strength on Titration Curves of the γ-Crystallins

Next, the effect of ionic strength on the γ-crystallin titration curves was studied. Titration experiments were performed with γ_{IV} in 0.01 M KCl and 0.1 M KCl. Both the acidic and basic ranges were investigated. The results of these experiments are presented in figures 4.14 and 4.15.

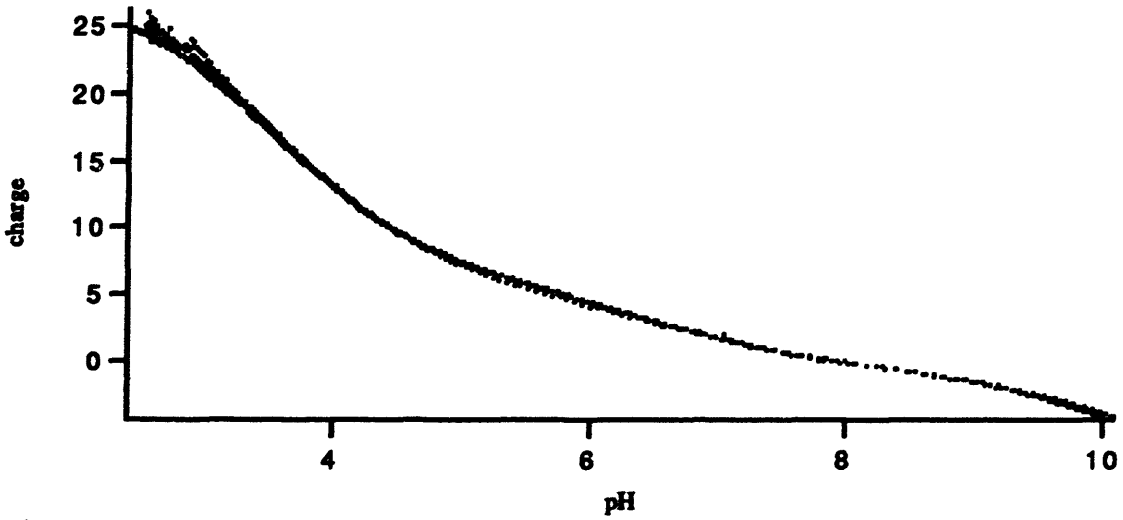


Figure 4.12 -- Three experiments are plotted: titration of γ_{IIIb} in 0.1 M KCl, 0.1 M NaCl, and 0.1 M NaBr. Since the curves are almost identical, points Corresponding to different conditions were not separately indicated.

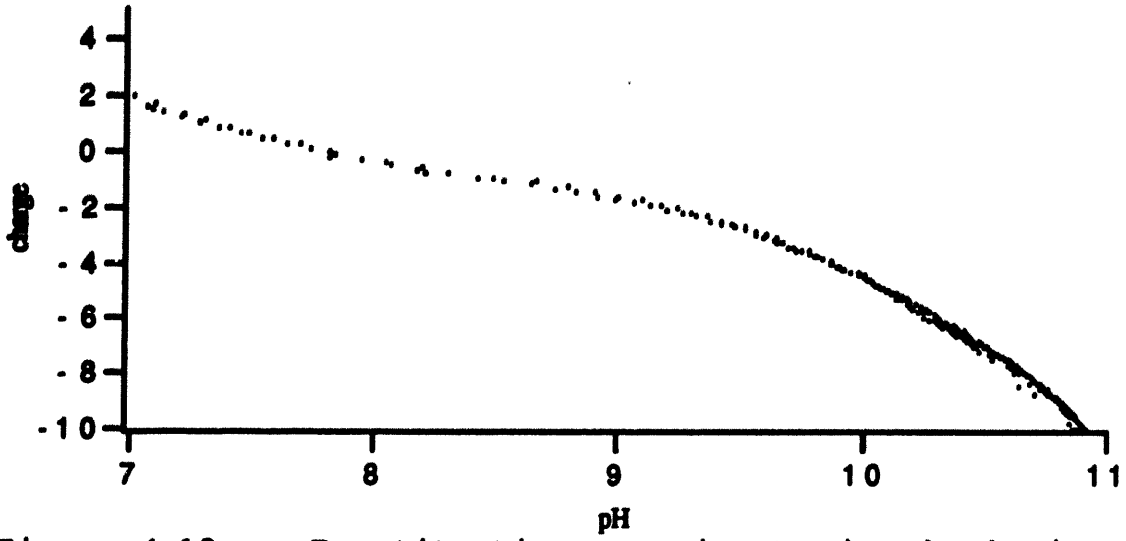


Figure 4.13 -- Two titration experiments in the basic range with γ_{IV}. The ionic strength was 0.1 M in both, but the ionic identity was changed from KCl in one experiment to NaCl in the other. No effort was made to identify the points with their electrolyte since the curves overlie one another.

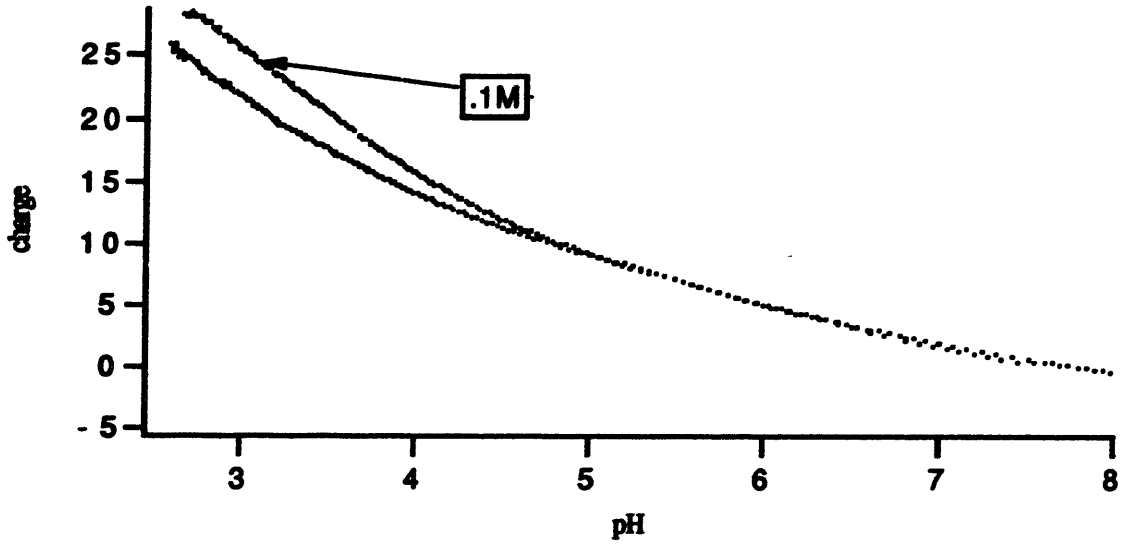


Figure 4.14 -- Titration of γ_{IV} in 0.1 M KCl and in 0.01 M KCl in the acidic range.

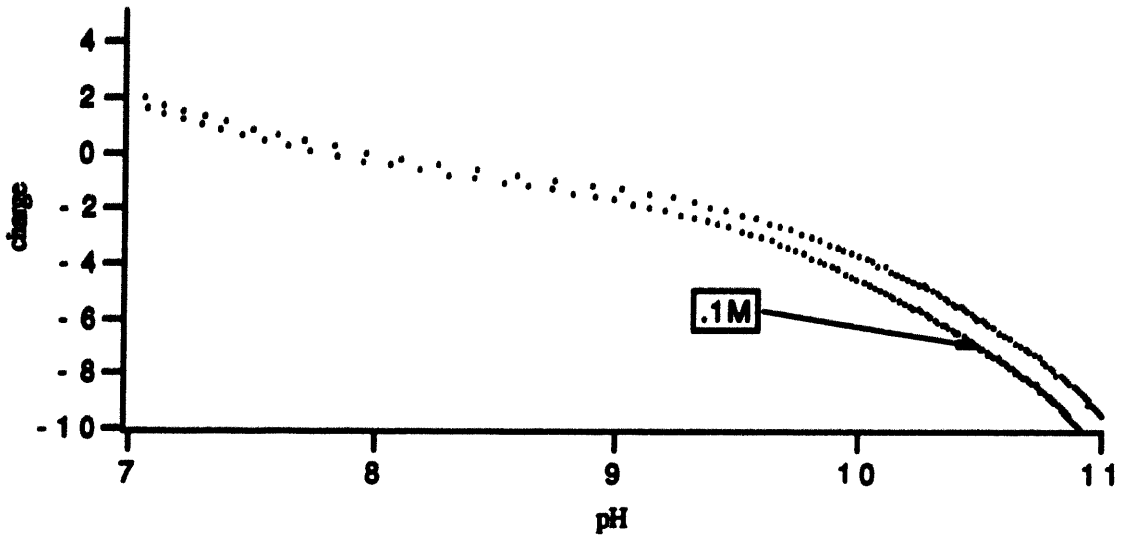


Figure 4.15 -- Basic range titration of γ_{IV} in 0.1 M KCl and 0.01 M KCl.

Qualitatively, one would expect that decreasing ionic strength would increase the strength of intramolecular electrostatic interactions by reducing screening. This would result in the more charged configurations of the protein becoming less energetically favorable. As screening is reduced, the work involved in increasing the charge by one unit increases due to the increased strength of the electrostatic repulsion. This additional work can be thought of as modifying the proton binding energy, and hence effecting the *pK*s (see section 2.8). Above the isoelectric point of the protein (≈ 8 for the γ-crystallins) the net protein charge will be negative. In the case of basic *pK*s, where the positively charged residue becomes neutral with the loss of a proton, the *pK* will shift to higher values. Since the majority of residues with *pK*s above the isoelectric point

are basic, the effect of decreased screening will be, on average, to flatten the curve above the isoelectric point. This corresponds to the basic *pKs* being shifted to higher pH values. Likewise, the effect of decreased screening below the isoelectric point, where the protein is positively charged, is to reduce the *pKs* of acidic residues. This is the case because the net positive charge of the protein interacts favorably with the negative charge of the deprotonated acidic residue. Because the majority of residues with *pKs* below the isoelectric point are acidic, the net effect of decreased screening is a flattening of the titration curve below the isoelectric point as well. On average, decreasing screening would cause *pKs* to be shifted away from the middle *pH* range, causing the protein to carry less net charge. As shown in figures 4.14 and 4.15, this was exactly the effect observed.

Chapter 5 - The Effect of Electrostatic Interactions on Protein *pK*s and the Electrostatic Protein-Solvent Interaction Energy

5.0 Introductory Remarks

This chapter attempts to answer two important questions concerning electrostatic interactions in the γ-crystallins: (1) Can the γ-crystallins' acid-base titration curves be understood in terms of electrostatic corrections to intrinsic proton binding energies? (2) What is the electrostatic contribution to the protein-solvent interaction energy for the γ-crystallins? In order to perform these calculations it is necessary to calculate the work required to place charge on the protein molecules both in solution and in vacuum.

Two methods are used to theoretically determine the titration curves. In the first method (the Linderstrom-Lang model) the protein is treated as a spherically symmetric shell of charge, and the electrolyte is described by using the linearized Poisson-Boltzmann equation. The shortcomings of this technique are that it does not associate *pK* changes

with specific charge sites in the protein, and that the Linearization of the Poisson-Boltzmann equation is not a legitimate apporoximation except near the isoelectric point of the protein. The failure of the linearized Poisson-Boltzmann equation is corrected for by using the unlinearized equation. In order to locate charge within the protein, another method of determining effective *pK*s is used. In this method (the Kirkwood-Tanford model), electrostatic corrections to proton binding energies are calculated by treating charged sites within the protein as point charges. The locations of these point charges are determined by x-ray crystallographic studies. This approach has the benefit of predicting the exact location of charge as a function of *pH*. It requires detailed structural information, however, which was only available for γ_{II}.[22]

The electrostatic interaction energy between protein and solvent is an important parameter in theoretically understanding the phase separation process. In order to determine this interaction energy it is necessary to calculate the difference between the work of charging the protein in solution and in vacuum. This is done as a function of *pH* for all of the γ-crystallins.

Before answering the questions posed above some general aspects of modeling protein electrostatics are discussed. These include a discussion of the validity of treating

[22]Wistow, G. *J. Mol. Bio.* **1983**, 170, 175.

protein molecules as uniform dielectrics, and a review of the Poisson-Boltzmann method for describing electrolytic solvents. The specific models used to describe the proteins are reviewed in detail.[23,24,25]

5.1 General Considerations in Modeling Protein Electrostatics

5.1.1 Treating Proteins as Uniform Dielectrics

The dielectric constant of proteins in bulk has been measured over a broad frequency range (0 to 20 GHz) in various states of hydration, ranging from lyophilized to in solution. In each instance, the dielectric constant has been determined to lie between 2 and 4.[26,27] A slight frequency dependence is observed in the megahertz range, and is believed to be the result of rotating dipolar side-chains at the surface.[28] The dielectric constant of protein can also be calculated using the Clausius-Mosotti theory.[29] This gives results consistent with experiment.

---

[23]Linderstrom-Lang, K. *Trav. Lab. Carlsberg* **1924**, 15, 7.
[24]Tanford, C.; Kirkwood, J. G. *J. Am. Chem. Soc.* **1957**, 79, 20.
[25]Sharp, K.; Honig, B. *Chem. Scrip.* **1988**, 29A, 71.
[26]Takashima, S.; Schwan, H. P. *Phys. Chem.* **1965**, 69, 4176.
[27]Harvey, S. C.; Hoekstra, P. *J. Phys. Chem.* **1972**, 76, 2987.
[28]Pennock, B. D.; Schwan, H. P. *J. Phys. Chem.* **1969**, 73, 2600.
[29]Reitz, J. R.; Milford, F. J. *Foundations of Electromagnetic Theory;* Addison-Wesley: Reading, 1967.

These statements do not resolve whether or not treating a protein molecule as a uniform dielectric is legitimate for the purpose of describing electrostatic interactions between charged sites within the protein. The dielectric constant depends on the polarizability of the atoms which make up the dielectric material. Microscopically, the dielectric constant fluctuates on atomic distance scales depending on the exact location and identity of atoms and molecules making up the dielectric material. The macroscopic dielectric constant represents a spatial average over these fluctuations. In order for the notion of a macroscopic dielectric constant to make sense, the distance scales over which interactions are being investigated must be large compared to atomic distance scales. The case of charge interactions within a protein can come close to violating this constraint. The polarizability varies on a scale of atomic sizes. The majority of interactions between charged residues within a protein are over distance scales large compared to this. However, some significant interactions arise from charge pairs which are separated by only a few angstroms. Over this scale the significance of the macroscopically determined dielectric constant becomes uncertain.[30] Despite this, the standard models for describing protein electrostatics make this continuum assumption. In doing so, these models reproduce experimental results with

[30]Harvey, S. *Prot.* **1989**, 5, 78.

reasonable accuracy. Some attempts to include the spatial variation in the atomic polarizabilities of individual atoms have been undertaken.[31] Although this approach is promising, it is not practical for the calculations described below.

5.1.2 Protein Shape and Size

The γ-crystallins are a family of highly homologous globular proteins found in the fiber cell cytoplasm of the bovine ocular lens.[32,33,34] All of these proteins have a molecular weight of approximately 21 kilodaltons.[35,36] The most extensively studied crystallin is γ_{II} which has had its structure determined to within 1.9 angstroms by x-ray crystallography.[37] On the basis of these studies a spherical approximation to the protein structure is reasonable, although in reality the proteins are slightly ellipsoidal.

Having chosen to model the protein molecules as dielectric spheres the next question involves choosing the appropriate radius. Folded protein generally includes some

---

[31]Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, 103, 227.
[32]Breitman, M. L. *Proc. Natl. Acad. Sci.* **1984**, 98, 7762.
[33]Bjork, I. *Exp. Eye Res.* **1961**, 1, 145.
[34]Meakin, S. O.; Breitman, M. L.; Tsui, L. C. *Mol. Cell. Biol.* **1985**, 5, 1408.
[35]Harding, J. J.; Dilley, K. J. *Exp. Eye Res.* **1976**, 22, 1.
[36]Bindels, J. G.; et. al. *Comp. Biochem. Physiol. B* **1983** , 76, 47.
[37]Wistow, G. *J. Mol. Bio.* **1983**, 170, 175.

tightly bound solvent molecules which must be included in a theoretical determination of the protein's size. The effective radius of a protein must be chosen so that the corresponding volume includes bound water molecules, i.e.,

$$R = \left(\frac{3}{4\pi}\ (V_p + N_w V_w)\right)^{\frac{1}{3}} \tag{5.1}$$

where V_p is the protein volume, N_w is the number of bound molecules of water per protein, and V_w is the volume of a water molecule. This can be expressed in terms of the specific volumes (inverse mass density) of the protein and water in bulk, $\mathcal{V}_p$ and $\mathcal{V}_w$. In these terms the effective volume of a protein molecule with bound water can be written as,

$$V_p + N_w V_w = m_p \mathcal{V}_p + N_w m_w \mathcal{V}_w \tag{5.2}$$

where m_p and m_w are the mass of a protein and molecule and water molecule respectively. This can be rewritten as,

$$V_p + N_w V_w = m_p \left(\mathcal{V}_p + \frac{m_w}{m_p} N_w \mathcal{V}_w\right) = \frac{M}{A}\ (\mathcal{V}_p + f\mathcal{V}_w) \tag{5.3}$$

where M is the molecular weight of the protein, A is Avogadro's number, and f is the amount of bound solvent in grams per gram of dry protein. This allows the effective

protein radius to be written as[38],

$$R = \left(\frac{3M}{4\pi A}\ (V_p + fV_w)\right)^{\frac{1}{3}} \qquad (5.4)$$

Using this method, 18 Å is the equivalent radius for the γ-crystallins.

5.1.3 Description of the Aqueous Electrolytic Solvent: The Poisson-Boltzmann Electrolyte

Treating the protein interior as uniform dielectric enables us to solve the Poisson equation and determine the general form of the electrostatic potential within the protein. This will be done for specific charge distributions in sections to follow. In order to determine the specific solution however, it is necessary to solve for the general solution of the electrostatic potential in the solvent outside of the protein and match this solution at the boundary. The solvent is also an isotropic dielectric, however the solvent also contains dissolved small ions. The presence of these small ions introduces a charge density in the Poisson equation outside of the protein. This section describes the Poisson equation in a solvent where there are

[38]Tanford, C. *The Physical Chemistry of Macromolecules;* Wiley: New York, 1961.

dissolved ions.[39] The resulting equation for the electrostatic potential is referred to as the Poisson-Boltzmann equation. After deriving the Poisson-Boltzmann equation, the case of a spherically symmetric charge distribution will be studied in detail, with an analytic solution for the electrostatic potential being the result. The results of this section are used in later sections where the electrostatic potential and various interaction energies are determined for specific charge distributions within proteins.

The starting point for the Poisson-Boltzmann description of an electrolyte lies, as the name suggests, in two equations: one from electrostatics, the Poisson equation, and the other from statistical mechanics, the Boltzmann equation, i.e.,

$$\varepsilon \nabla \cdot \mathbf{E}(\mathbf{x}) = -\varepsilon \nabla^2 \phi(\mathbf{x}) = 4\pi \rho(\mathbf{x}) \tag{5.5}$$

and

$$n_i(\mathbf{x}) = n_i^{o}\, e^{-\beta z_i q \phi(\mathbf{x})} \tag{5.6}$$

where ε is the dielectric constant ($\approx$80 for water), $\mathbf{E}$ is electric field, ϕ is the electrostatic potential, ρ is the free charge density, the subsript i refers to the species of ion, n_i is the number density, n_i^{o} is the average number

---

[39]Bockris, J. O.; Reddy, A. K. N. *Modern Electrochemistry;* Plenum Press: New York, 1970.

density, z_i is the valence, $\beta = 1/kT$, and q is the fundamental charge. The charge density can be recast in terms of the Boltzmann equation as,

$$\rho(x) = q \sum_i n_i(x)\ z_i = q \sum_i n_i^{\circ}\ e^{-\beta z_i q\phi(x)}\ z_i \qquad (5.7)$$

Considering an electrolyte in which there is only one type each of anion and cation and both anion and cation have valence one (a symmetric 1-1 electrolyte), this can be written as,

$$\rho(x) = q\ n^{\circ}\ (\ e^{-q\phi\beta} - e^{q\phi\beta}\ ) \qquad (5.8)$$

where n° is the ionic strength of the solution.

Next, the Poisson equation is used with equation 5.8 to obtain the Poisson-Boltzmann equation,

$$\nabla^2\phi(x) = -\frac{4\pi}{\varepsilon}\ q\ n^{\circ}\ (e^{-q\phi\beta} - e^{q\phi\beta}) \qquad (5.9)$$

The solution to this equation which satisfies the boundary conditions provides the electrostatic potential and hence the charge distribution of electrolyte in solution. This equation, however, can be simplified further.

By considering the regime where $q\phi\beta << 1$, and by

expanding the exponentials in equation 5.9 to first order, the Poisson-Boltzmann equation can be linearized. The equation obtained from this procedure is,

$$\nabla^2\phi(x) = \frac{8\pi q^2 n^o}{\varepsilon kT}\phi(x) \qquad (5.10)$$

The linearized Poisson-Boltzmann equation adequately describes many features of electrolytic solutions, and it will be used below. Its shortcommings will however be evident in section 5.2.1, and modifications will be discussed in section 5.2.2.

5.1.3.1 Solution of Poisson-Boltzmann Equation for Spherically Symmetric Charge Distributions

In this section the linearized Poisson-Boltzmann equation is solved for the case of a spherically symmetric charge distribution. This solution will be used in the Linderstrom-Lang model of protein charge, described in section 5.2.1.

Considering a spherically symmetric charge distribution (for example a uniformly charged spherical shell) in a symmetric 1-1 electrolyte with dielectric constant ε, the linearized Poisson-Boltzmann equation takes the form,

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\phi(r)}{dr}\right) = \frac{8\pi q^2 n^o}{\varepsilon kT}\phi(r) \qquad (5.11)$$

where r is the radial distance from the center of the charged shell, and other symbols are as defined previously. The constants on the right hand side can be collected together and defined as,

$$\kappa^2 = \frac{8\pi}{\varepsilon kT} n^o q^2 \qquad (5.12)$$

Defining the function μ in terms of ϕ by, $\mu(r) = r\phi(r)$, equation 5.11 becomes,

$$\frac{d^2\mu}{dr^2} = \kappa^2 \mu \qquad (5.13)$$

The solution to this equation is recognized as,

$$\mu(r) = Ae^{-\kappa r} + Be^{\kappa r} \qquad (5.14)$$

Since $\mu \rightarrow 0$ as $r \rightarrow \infty$, B must be equal to zero, leaving,

$$\mu(r) = Ae^{-\kappa r} \qquad (5.15)$$

or, finally,

$$\phi(r) = \frac{Ae^{-\kappa r}}{r} \qquad (5.16)$$

Kappa, defined in equation 5.12, can be identified as an inverse screening length. In other words, the ions arrange themselves in such a way as to cause screening of the field from the central charge. This screening occurs over a distance characterized by the length κ^{-1}. The constant, A, in equation 5.16 is determined by the boundary conditions relevant to a specific problem.

5.1.3.2 The Ion Exclusion Region

In section 5.1.3.1 ions in solution were treated as point charges. The Poisson-Boltzmann description of an electrolyte can be made more accurate by imposing an ion exclusion region around the surface of the object being studied. A natural choice for this distance is the radius of the counter-ions. The size of the like charged ions is less significant since the likelihood of encountering a like charged ion in close proximity to the macroion surface is small. In the case of proteins, however, there can be both types of charge on the protein surface. In this instance, the appropriate choice for the size of the ion exclusion region is a weighted average of the radii of the ions in

solution. This simple technique of imposing an ionic exclusion region renders the Poisson-Boltzmann theory of electrolytic solutions an accurate, computationally economical manner for describing ionic solutions near macroions and charged surfaces.

5.2 Specific Continuum Electrostatic Models Used to Study Proteins

Two methods of describing electrostatic interactions between charge on a protein will be discussed in this section. It is the purpose of both of these methods to provide a solution for the electrostatic potential in and around charged protein molecules in solution. The first method, attributable to Linderstrom-Lang, treats the protein molecule as a uniformly charged spherical shell. The model of Kirkwood and Tanford is considered next. This model treats the protein charge distribution as a sum of point charges whose positions are determined by x-ray crystallographic structural data. The spherical approximation to protein shape is preserved. In both of these techniques the Poisson equation must be solved for all of space. Except in the ion exclusion region, the solvent is described by the Poisson-Boltzmann equation.

5.2.1 The Linderstrom-Lang Model and the γ-Crystallins: Work of Charging, Proton Binding, and Electrostatic Protein-Solvent Interaction Energy

The Linderstrom-Lang model treats the protein as a spherical dielectric cavity of uniform dielectric constant.[40] This spherical dielectric region is considered to be in solvent which obeys the linearized Poisson-Boltzmann equation.

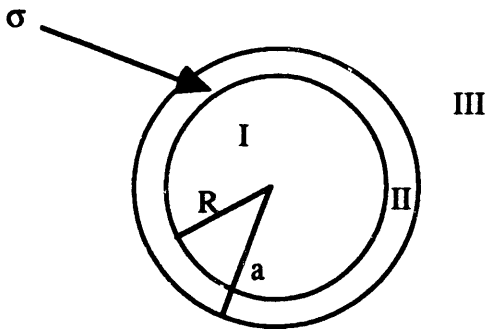


Figure 5.1 -- Spherical protein with surface charge σ in the Linderstrom-Lang model of protein charge. Region I, the protein interior, is a uniform dielectric with dielectric constant of 4. Region II is the ion exclusion region, in which the dielectric constant is the same as in region III. Region III is the solvent, treated as a Poisson-Boltzmann electrolyte with dielectric constant of 78.

The finite size of ions in solution is accounted for in the Linderstrom-Lang model by an ion exclusion radius as

[40]Linderstrom-Lang, K. *Trav. Lab. Carlsberg* **1924**, 15, 7.

described in section 5.1.3.2. The choice of radius for the protein is discussed in section 5.1.2. For the γ-crystallins the radius is taken to be 18 Å. The protein charge in this model is taken to be uniformly smeared over the surface of the spherical protein. See figure 5.1.

To determine the electrostatic potential in the Linderstrom-Lang model it is necessary to solve the Poisson equation in each of the three regions demonstrated in figure 5.1. The solutions must then be matched at the boundaries. Regions I and II represent the interior of the protein and the ion exclusion regions respectively. Since there is no charge within these regions, they both must satisfy the Laplace equation. There is, however, a surface charge ($\sigma=Q/4\pi R^2$) at the boundary between these two regions. Surface charge distributions do not cause discontinuity of the electrostatic potential. They do however cause a discontinuity of the electric field and this effect is taken into account in the boundary conditions of equation 5.21. Region III, the solvent region, satisfies the Poisson-Boltzmann equation as previously described. The solutions to these equations for cases with spherical symmetry can be written as,

$$\phi_I(r) = C_1 \qquad r < R \qquad (5.17)$$

$$\phi_{II}(r) = C_2 + \frac{C_3}{r} \qquad R < r < a \qquad (5.18)$$

$$\phi_{III}(r) = C_4 \frac{e^{-\kappa r}}{r} \qquad r > a \qquad (5.19)$$

These solutions must match at the boundaries, i.e.,

$$\phi_I\ (r) = \phi_{II}(r) \qquad (5.20a)$$

$$\phi_{II}\ (a) = \phi_{III}(a) \qquad (5.20b)$$

and,

$$\varepsilon_I\ [\frac{\partial \phi_I}{\partial n}]_R = \varepsilon_{II}\ [\frac{\partial \phi_{II}}{\partial n}]_R \qquad (5.21a)$$

$$\varepsilon_{II}\ [\frac{\partial \phi_{II}}{\partial n}]_a = \varepsilon_{III}\ [\frac{\partial \phi_{III}}{\partial n}]_a \qquad (5.21a)$$

Noting that the dielectric constants in regions II and III are equal, the constants in equations 5.17, 5.18, and 5.19 can be determined, providing the electrostatic potential in all three regions,

$$\phi_I(r) = \frac{Q}{\varepsilon R}\ (1 - \frac{\kappa R}{1+\kappa a}) \qquad r < R \qquad (5.22)$$

$$\phi_{II}(r) = \frac{Q}{\varepsilon r}\left(1 - \frac{\kappa r}{1+\kappa a}\right) \qquad R < r < a \qquad (5.23)$$

$$\phi_{III}(r) = \frac{Q}{\varepsilon(1+\kappa a)}\,\frac{e^{-\kappa(r-a)}}{r} \qquad r > a \qquad (5.24)$$

where κ is as defined in equation 5.12, Q is the net protein charge, R is the protein radius, a is the ion exclusion radius, and ε is the dielectric constant of the solvent, i.e. 78.

5.2.1.1 Work of Charging in the Linderstrom-Lang Model

In the Linderstrom Lang model, it is necessary to determine the work of charging a protein in order to correct the titration curves for electrostatic interactions and in order to determine the electrostatic protein solvent interaction energy. For this reason the work of charging is discussed first.

The work of charging a protein molecule is the work required to place the protein charge on the protein. The infinitesimal work, dW_{el}, that is required to infinitesimally increase the charge on the protein is,

$$dW_{el} = \phi(q)\, dq \qquad (5.25)$$

where ϕ is the electrostatic potential at the site of the infinitesimal charge increase, q is charge. The potential is a function of both solvent environment and protein charge, and is given in equation in equations 5.22 -5.24. The total work required to charge a protein can be written in terms of equation 5.25 as,

$$W_{el} = \int_0^Q \phi(q,R)\ dq \tag{5.26}$$

where Q is the final charge. The electrostatic potential determined above can be used in equation 5.26 to determine the work explicitly,

$$W_{el} = \int_0^Q \phi(q, r=R)\, dq = \frac{Q^2}{2\varepsilon R}\left(1 - \frac{\kappa R}{1+\kappa a}\right) \tag{5.27}$$

This expression was evaluated as a function of pH for the γ-crystallins. To do this, the charge versus pH results of chapter 4 were used. The results are plotted in figures 5.2 - 5.9. The radius of the protein was taken to be 18 $\AA$, the ion exclusion radius 20 $\AA$, and κ^{-1}=7 $\AA$.

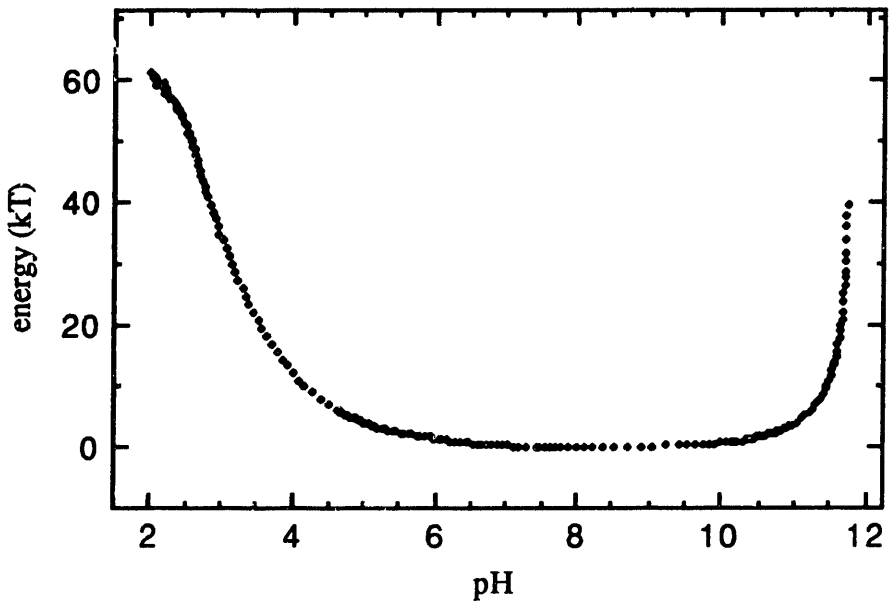


Figure 5.2 -- Linderstrom-Lang work of charging γ_{II} in 0.1 M KCl.

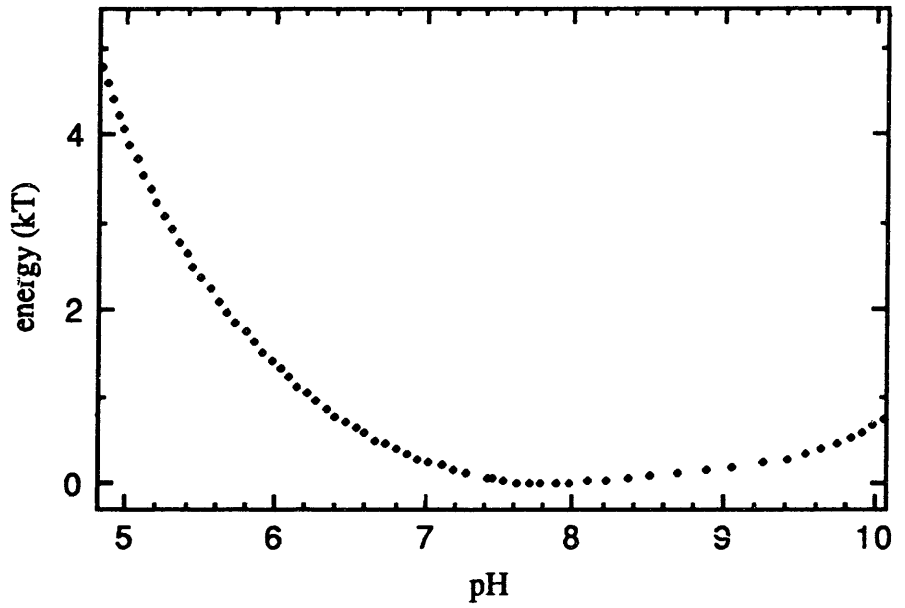


Figure 5.3 -- Linderstrom-Lang work of charging γ_{II} in 0.1 M KCl. This is an enlargement of figure 5.3 to illustrate behavior near physiologic *pH*.

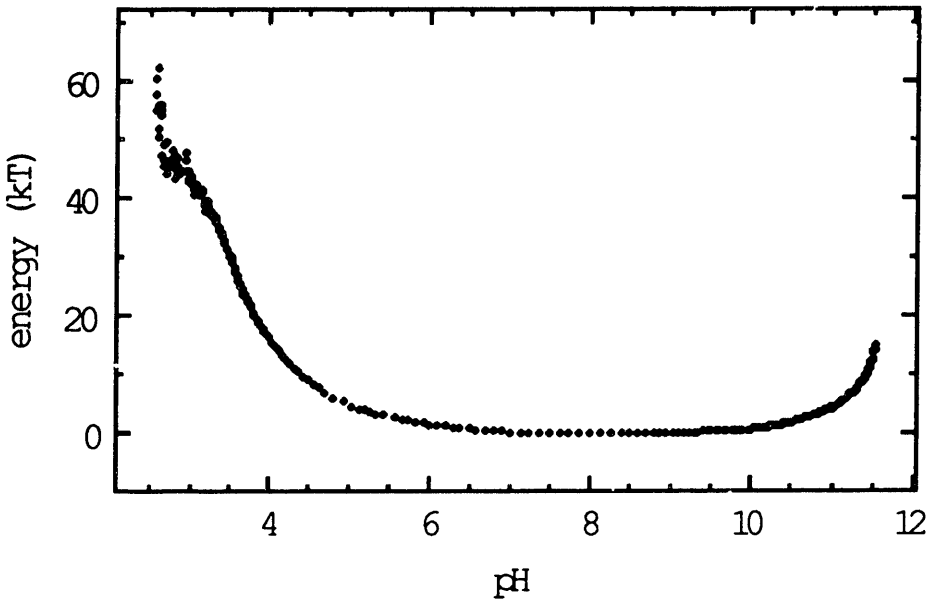


Figure 5.4 -- Linderstrom-Lang work of charging γ_{IIIa} in 0.1 M KCl.

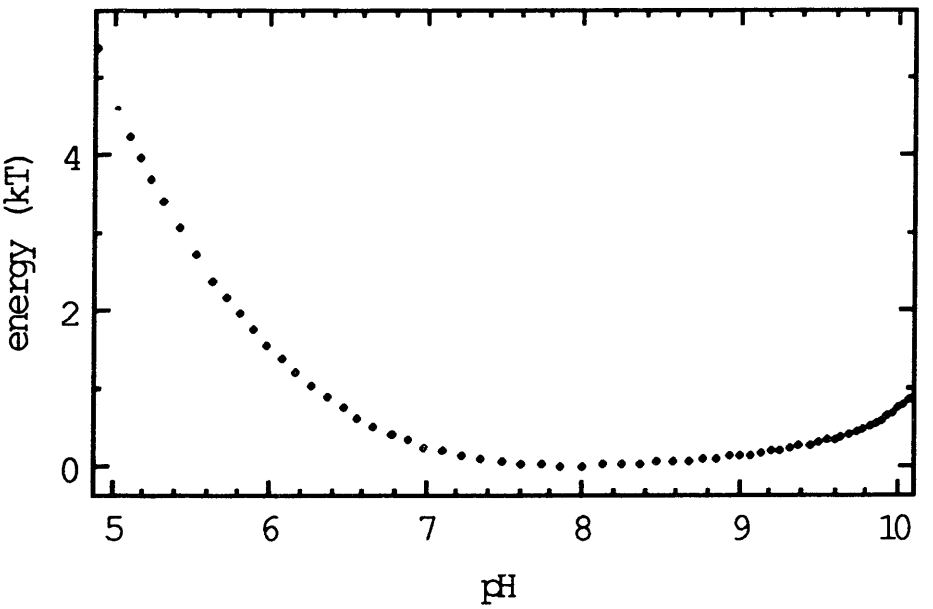


Figure 5.5 -- Linderstrom-Lang prediction for the work of charging γ_{IIIa} in 0.1 M KCl. This is an enlargement of figure 5.5 to illustrate behavior near physiologic *pH*.

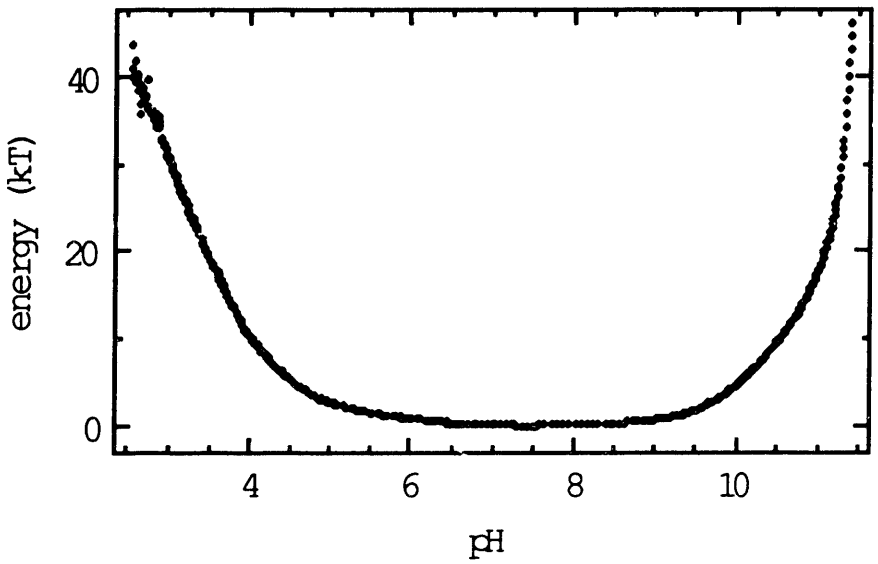


Figure 5.6 -- Linderstrom-Lang prediction for the work of charging γ_{IIIb} in 0.1 M KCl.

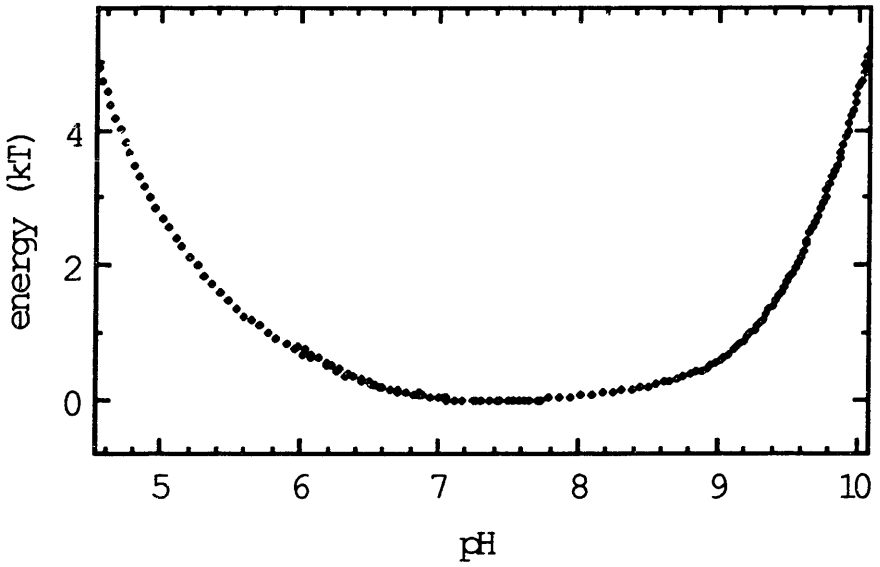


Figure 5.7 -- Linderstrom-Lang prediction for the work of charging γ_{IIIb} in 0.1 M KCl. This is an enlargement of figure 5.7 to illustrate behavior near physiologic *pH*.

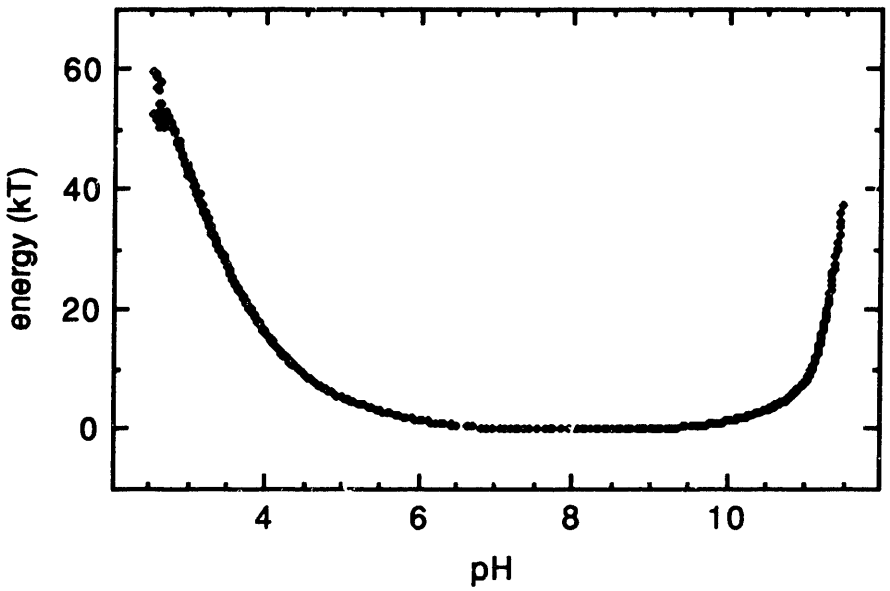


Figure 5.8 -- Linderstrom-Lang work of charging γ_{IV} in 0.1 M KCl.

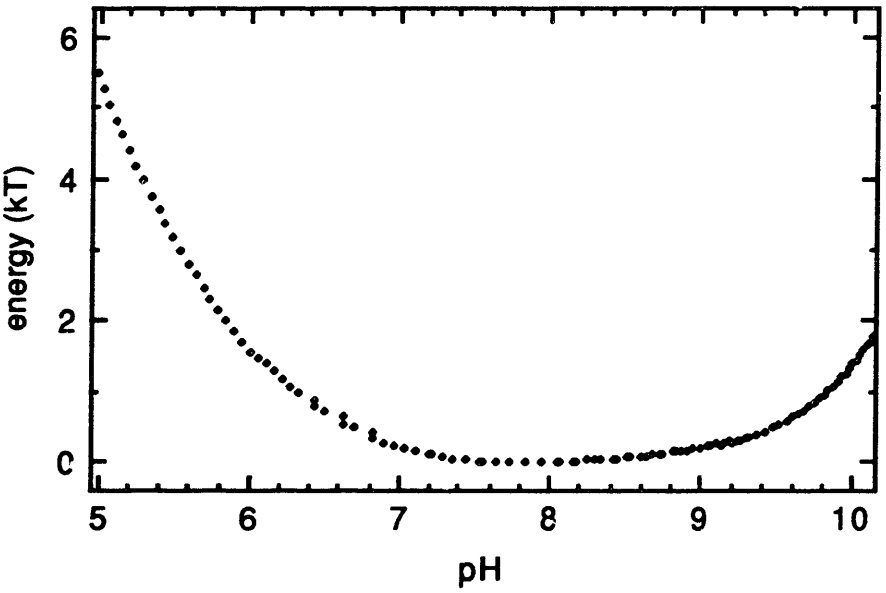


Figure 5.9 -- Linderstrom-Lang prediction for the work of charging γ_{IV} in 0.1 M KCl. This is an enlargement of figure 5.9 to illustrate behavior near physiologic *pH*.

5.2.1.2 Electrostatic Effects on Acid-Base Titration in the Linderstrom-Lang Model

In this section, the Linderstrom-Lang model will be used to predict the *pK* shifts (i.e. changes from the intrinsic *pK* values) which result from electrostatic interactions between the titrateable sites within a protein and between titrateable sites and solvent. Recall from section 2.8 that the apparent *pK* of a titrateable residue is related to the intrinsic *pK* (pK_o) and the electrostatic correction to the proton binding energy of the titrateable site, *W*, by the equation,

$$pK = pK_o + \Delta pK \tag{5.28}$$

where,

$$\Delta pK = \frac{-W}{kT\ ln(10)} \tag{5.29}$$

The total proton binding energy can be thought of as the work of taking a proton from far away and placing it at the titrateable site. This work can be subdivided into the intrinsic binding energy of the titrateable site in the absence of other interactions, and the work due to the

interaction between the proton charge and other charge within the protein. This latter work, W, can be expressed in terms of the work of charging the protein, W_{el}., as,

$$W = W_{el}(Z+1) - W_{el}(Z) = (1)\ \frac{\partial W_{el}}{\partial Z} \tag{5.30}$$

where Z is the net charge in units of fundamental charge. Using the results of section 5.2.1.1 for W_{el}, equation 5.30 becomes,

$$W = \frac{Zq^2}{\varepsilon R}\left(1 - \frac{\kappa R}{1+\kappa a}\right) \tag{5.31}$$

where Z is the number of charges on the protein, R is the radius of the protein, a is the ion exclusion radius, q is the fundamental charge, and κ is the inverse screening length as described in equation 5.12. For ease of writing, the following definition is made,

$$w \equiv \frac{W}{Z}\ \frac{1}{kT\ ln(10)} \tag{5.32}$$

Referring to section 2.7, the number of protons binding a protein in the absence of interactions can be written as,

$$\langle N \rangle = \sum_{k=1}^{kmax} \frac{n_k}{1 + 10^{pH-pK_{ko}}} \tag{5.33}$$

where N is the number of protons bound, n_k is the number of residues in each class, and pK_{ko} is the pK associated with each type of residue. Refer to table 5.1 for a list of the residue and intrinsic pKs for the case of γ_{II}. The net charge, in units of fundamental charge, can be written as $Z = N - Z_o$, where Z_o is the number of acidic residues (i.e. those with charge between -1 and 0). Using equation 5.33, the net charge on a protein can be written as,

$$Z = \sum_{k=1}^{kmax} \frac{n_k}{1 + 10^{pH-pK_{ko}}} - Z_o \qquad (5.34)$$

where again, interactions have not been included. The pKs in equation 5.34 can be corrected, using the Linderstrom-Lang model, with equation 5.29, giving, $pK = pK_o - wZ$, where w is defined in equation 5.32. Using these corrected pKs, the titration curve of equation 5.34 becomes,

$$Z = \sum_{k=1}^{kmax} \frac{n_k}{1 + 10^{pH-pK_{ok}+wZ}} - Z_o \qquad (5.35)$$

Equation 5.35 is the predicted titration curve, including electrostatic corrections to pKs, in the Linderstrom-Lang model. Equation 5.35 can be numerically solved for the charge, Z, as a function of pH. This was done for γ_{II} and the results are presented in figure 5.10. The software to solve this equation was written in C and is

reproduced in appendix E. Unfortunately, the *pK*s of specific residues cannot be distinguished in this model. This is a significant shortcoming, not present in the Kirkwood-Tanford model, which is discussed below.

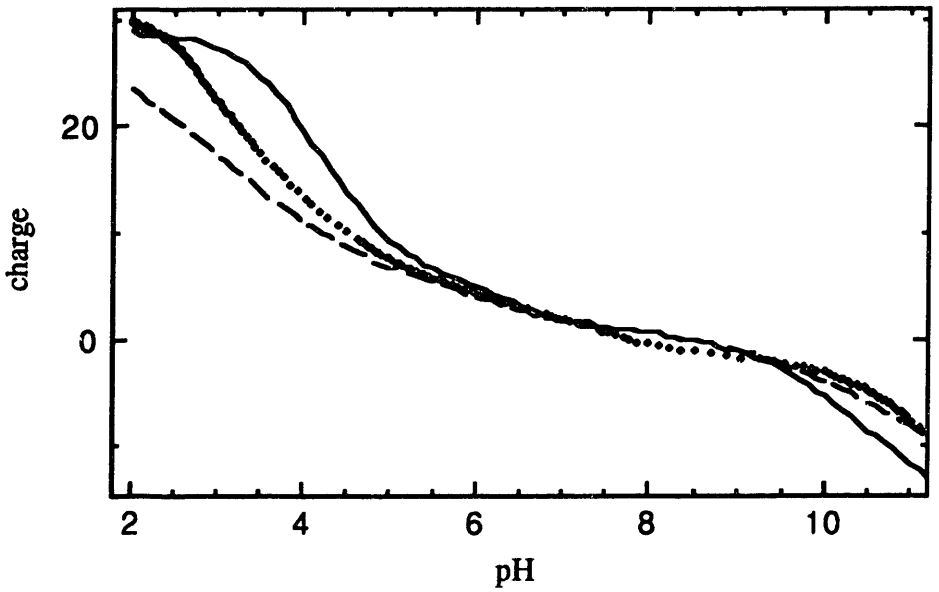


Figure 5.10 -- Theoretical Linderstrom-Lang titration curve for (dashed line) and theoretical curve neglecting electrostatic corrections (solid curve) compared with titration data for γ_{II}. The dots represent data.

The comparison between the theoretical Linderstrom-Lang titration curves and data in figures 5.10 through 5.12 illustrates a major shortcoming of the Linderstrom-Lang model. The model clearly fails to reproduce the data except at *pH* near the isoelectric point, where the net protein charge is near zero. Because the linearization of the Poisson-Boltzmann equation is not valid except near net charge of zero, the question arises as to whether using an

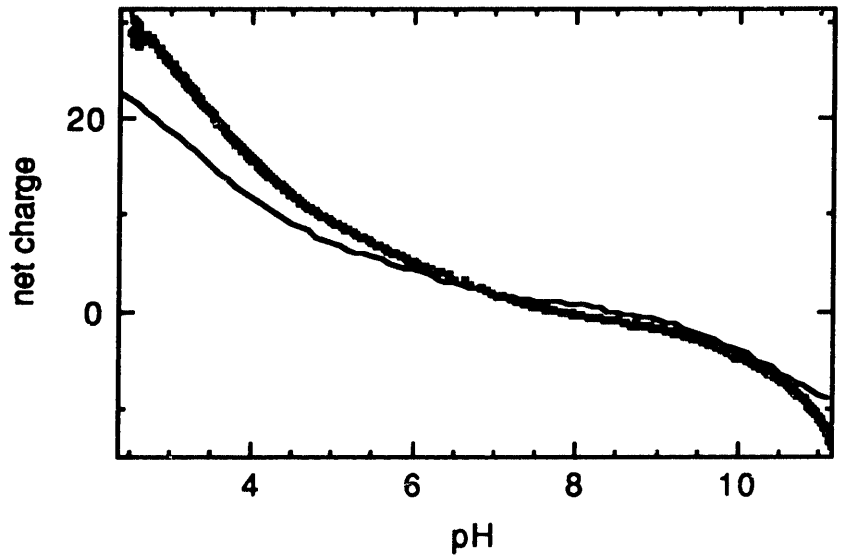
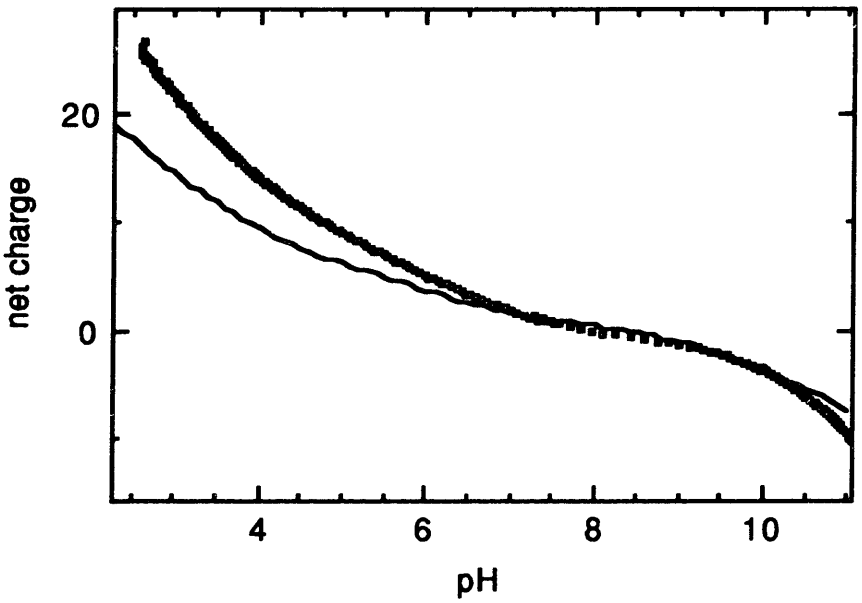


Figure 5.11 -- Comparison of theoretical Linderstrom-Lang titration curve (solid line) with data (dots) for γ_{IV} in 0.1 M KCl.



5.12 -- Comparison of theoretical Linderstrom-Lang titration curve (solid line) with data (dots) for γ_{IV} in 0.01M KCl. Corrections have been included for the change in ionic strength which results at *pH* extremes.

unlinearized version of the Poisson-Boltzmann equation would enable the smeared charge model to perform better.

5.2.1.3 The Electrostatic Interaction Between Protein and Solvent

The electrostatic interaction energy between a protein molecule and solvent is of interest because of its potential significance in determining the liquid-liquid phase boundary. This energy can be determined by calculating the work of charging the protein in solution and in vacuum. The difference between these works represents the electrostatic interaction energy between the solvent and the protein. The work of charging a protein in solution was determined in equation 5.27. The work of charging in vacuum can be obtained from 5.27 by setting $\kappa = 0$ and $\varepsilon = 1$. The resulting protein-solvent interaction can be written as,

$$E_{ps} = \frac{Q^2}{2\varepsilon_s R}\left(1 - \frac{\kappa R}{1+\kappa a}\right) - \frac{Q^2}{2R} \qquad (5.36)$$

where κ is the inverse screening length, a is the ion exclusion radius, R is protein radius, Q is protein charge, and ε_s is the solvent dielectric constant (78 for water).

The question as to whether or not the observed

dependence of the cloud point temperature, T_{cloud}, on protein charge can be easily understood in terms of this solvation energy can be addressed. Taratuta et. al. observed that, in lysozyme solutions, changing the *pH* so that the net protein charge increased by one unit resulted in a depression of T_{cloud} by approximately 5° C.[41]

From equation 5.36 the approximate change in solvation energy of lysozyme which results from changing the *pH* from 8 to 6 (changing the charge from about +7*e* to about +8*e*) can be determined and compared to the change in kT_{cloud} observed by Taratuta. Using this simple model, the change in solvation energy of a lysozyme molecule which results from changing the *pH* from 8 to 6 is about, $\Delta E_{ps} \approx -6$ eV. The change in kT_{cloud} which was observed to result from this change in *pH* is only $\Delta kT_{cloud} \approx 10^{-3}$ eV. This result suggests that the change in the interaction energy between protein and solvent does not alone adequately account for the observed dependence of T_{cloud} on charge.

The inability to predict changes in cloud point temperature in terms of the solvation energy alone is not unexpected. The relevant protein-solvent interaction energy is the difference in solvation energies between the protein-rich and protein-poor phases. Because the major contribution to the solvation energy arises from interactions with solvent

---

[41]Taratuta, V.; et. al. *J. Phys. Chem.* **1990**, 5, 2140.

that is very close to the protein surface, the protein-solvent interaction may be only weakly dependent on the concentration of protein. The role of protein-protein interactions, and their charge dependence, must also be considered. Answering these questions fully is beyond the scope of this thesis.

5.2.2 The Theoretical Titration Curve Using the Nonlinearized Modification of the Linderstrom-Lang Model

In order for the linearization of the Poisson-Boltzmann equation to be valid, the argument of the exponentials in equation 5.9 must be small, i.e.,

$$q\phi << kT \tag{5.37}$$

Except in the case of low net protein charge, this criteria is not fulfilled at the protein surface. For 0.1 M ionic strength and protein radius of 18 Å, $q\phi \approx (.003\text{eV})Z$, where Z is the net protein charge. Note that kT is .025 eV. At protein charge of 10, $q\phi$ is on the order of kT. This invalidates the Linearized Poisson-Boltzmann description of the electrolyte, and hence the Linderstrom-Lang model.

The problems associated with the validity of the linearized Poisson-Boltzmann equation can be avoided by simply not linearizing the Poisson-Boltzmann equation. The unlinearized Poisson-Boltzmann equation (equation 5.9) can not be solved analytically, even for spherically symmetric charge distributions. It has however been solved numerically and the results tabulated by Loeb, Overbeek, and Wiersema.[42] In order to determine the titration curves predicted by the nonlinear Poisson-Boltzmann equation it is only necessary to know the potential at the protein surface. Loeb's numerical work tabulates the surface potential of a spherical particle as a function of charge, size, and electrolyte conditions. It was found that the surface potential's dependence on protein charge was linear over the relevant range of parameters (see figure 5.13).

Making use of figure 5.13, the electrostatic potential at the protein surface was determined as a function of protein charge. In order to take the ion exclusion region into account, the sphere size was taken to be the ion exclusion radius and the potential difference was corrected to include the potential increase associated with crossing the ion exclusion region. This result was then used to modify the titration curve given by equation 5.34 the same manner that the Linderstrom-Lang corrections were imposed.

---

[42]Loeb, A. L.; Overbeek, J. G.; Wiersema, P. H. *The Electrical Double Layer Around a Spherical Colloid Particle* MIT Press: Cambridge, 1961.

The results of this analysis are presented in figures 5.14 through 5.16.

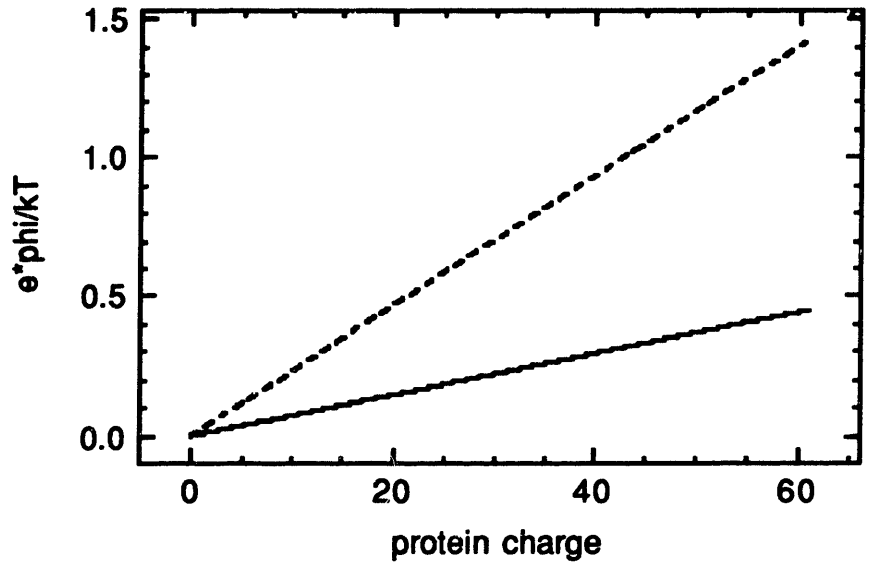
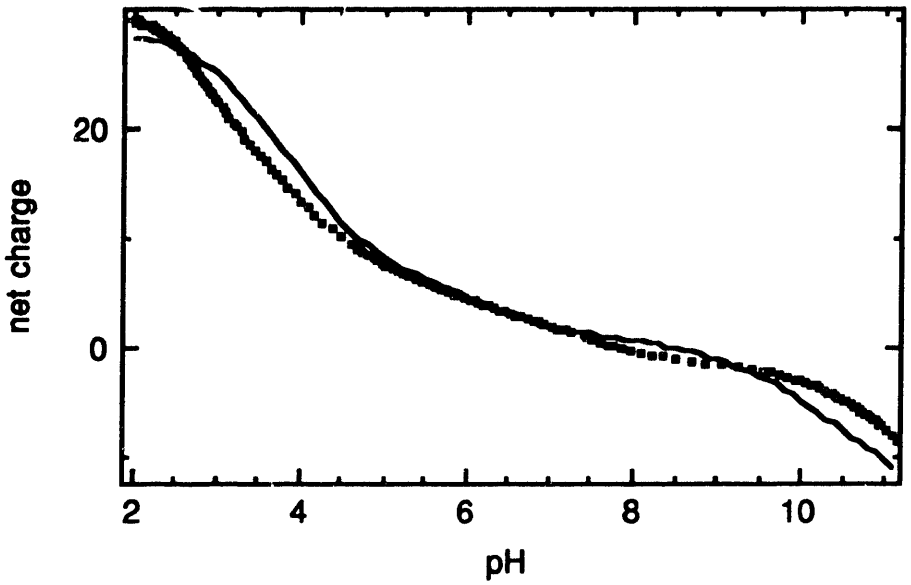


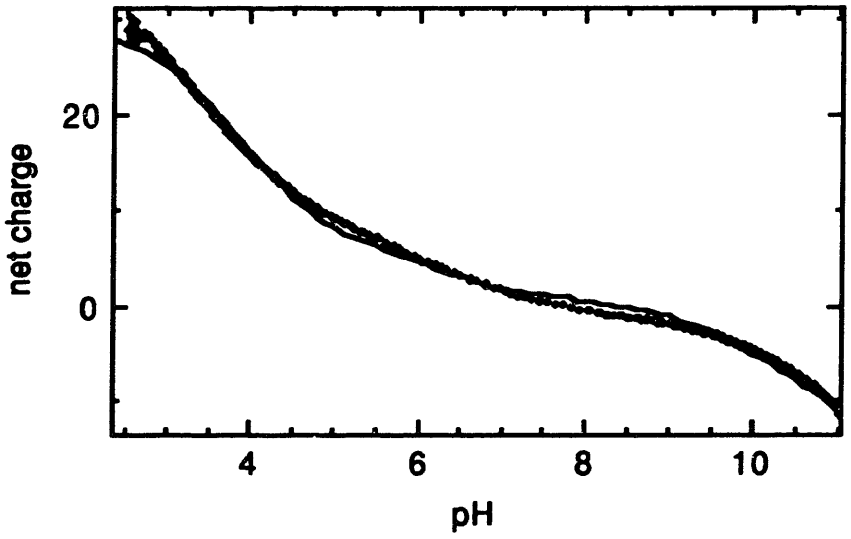
Figure 5.13 -- Electrostatic surface potential as determined numerically by Loeb et. al.[43] using the unlinearized Poisson-Boltzmann equation for a spherical charge distribution at 20 Å. The potential is presented in dimensionless form as a function of protein charge. The solid line represents results for 0.1 M ionic strength and the dashed line represents results for 0.01 M ionic strength. In order to determine the actual protein surface potential it is necessary to correct for the potential increase associated with crossing the ion exclusion region as described in the text.

---

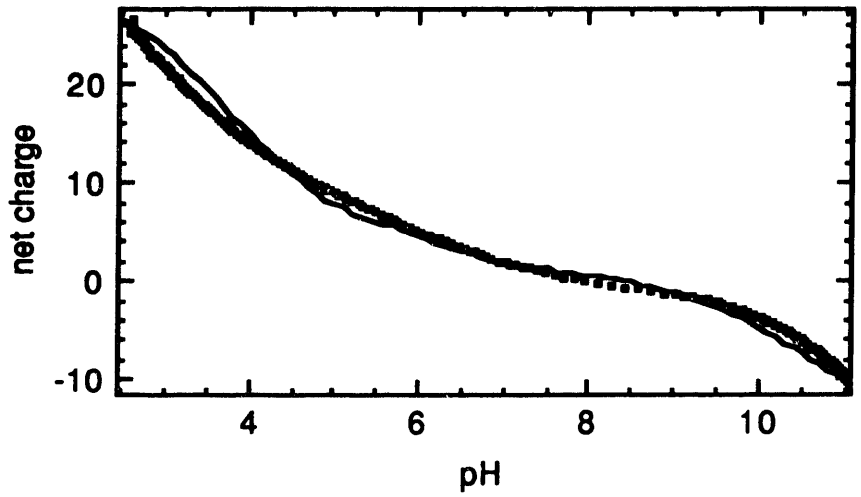
[43]Loeb, A. L.; Overbeek, J. G.; Wiersema, P. H. *The Electrical Double Layer Around a Spherical Colloid Particle* MIT Press: Cambridge, 1961.



Figureg 5.14 -- Theoretical titration curve (solid line) compared to data (dots) for γ_{II} in 0.1M KCl. The theoretical titration curve was obtained using the nonlinear Poisson-Boltzmann equation and treating the protein as a uniformly charged shell.



Figureg 5.15 -- Theoretical titration curve (solid line) compared to data (dots) for γ_{IV} in 0.1M KCl. The theoretical titration curve was obtained using the nonlinear Poisson-Boltzmann equation and treating the protein as a uniformly charged shell.



Figureg 5.16 -- Theoretical titration curve (solid line) compared to data (dots) for γ_{IV} in 0.01M KCl. The theoretical titration curve was obtained using the nonlinear Poisson-Boltzmann equation and treating the protein as a uniformly charged shell.

It is evident in figures 5.14 through 5.16 that the experimental titration curves are very well approximated by this method. The Linderstrom-Lang method failed to reproduce the experimental results as protein net charge was increased beyond about 10 units of fundamental charge. The nonlinear Poisson-Boltzmann modification to the Linderstrom-Lang model fits the data very well over the entire range for both 0.1 M and 0.01 M ionic strengths. Despite this model's success, it still does not assign *pK* corrections to specific residues within the protein. In order to accomplish this it is necessary to abandon the approximation of spherical symmetry of the charge distribution. The next section describes the

Kirkwood-Tanford model in which the protein charge is treated as point charge whose location within the protein is determined by x-ray crystallographic studies

5.2.3 The Kirkwood-Tanford Fixed Charge Model: Protein Charge as Point Charges with the Linearized Poisson-Boltzmann Solvent

The advantage of the Linderstrom-Lang model is that the location of charge sites within the protein need not be known. With the detailed structural information now available from x-ray crystallography the precise location of chargeable sites can, however, be determined. The Tanford-Kirkwood model takes advantage of this. By doing so, it refines the description of intraprotein charge interactions. Both the spherical approximation to protein shape and the linearized Poisson-Boltzmann description of the solvent are retained with an ion exclusion region near the surface. As in the Linderstrom-Lang model, both the region within and outside of the protein are treated as uniform dielectrics. In this model, however, titrateable groups are assigned a specific location within the sphere. Protein charge is treated as point charges at these sites.

As in the Linderstrom-Lang model, the objective is to

determine the electrostatic potential in and around the protein. Once this is done, the electrostatic potential can be used to determine the work of charging and corrections to the proton binding energies and *pK*s. The results are much more complicated than in the Linderstrom-Lang model, however, since there is no longer spherical symmetry in the charge distribution. See figure 5.17.

The problem is again divided into three regions as illustrated in figure 5.17. In region II the Laplace equation holds. The most general solution of the Laplace equation in spherical coordinates can be written as,

$$\phi_{II}(r,\theta,\psi) = \sum_{n=0}^{\infty}\sum_{m=-n}^{+n}\left(\frac{C_{nm}}{r^{n+1}} + G_{nm}r^{n}\right)P_n^{m}(\cos\theta)\ e^{im\psi} \qquad (5.38)$$

where the *C*s and *G*s are constants and *P* represents Legendre polynomials. The constants remain to be determined by boundary conditions.

In region I this solution must be modified in two respects. First, the constants C_{nm} must be zero to prevent a singularity at the origin. Second, the Green function appropriate to the collection of point charges within this region must be added, leaving,

$$\phi_{I}(r,\theta,\psi) = \sum_{n=0}^{\infty}\sum_{m=-n}^{+n}B_{nm}\ r^{n}\ P_n^{m}(\cos\theta)\ e^{im\psi} + \sum_{k}\frac{q_k}{\varepsilon_I|\boldsymbol{r}-\boldsymbol{r}_k|} \qquad (5.39)$$

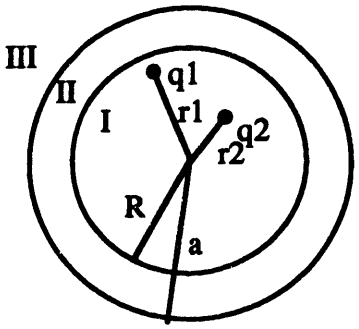


Figure 5.17 -- The Tanford-Kirkwood model in which charge is discretely located within a spherical protein. Region I is the protein, region II the ion exclusion zone, and region III the Poisson-Boltzmann solvent.

The general solution of the Poisson equation in region III is more complicated since the equation to be solved is not the Laplace equation, but rather the Poisson-Boltzmann equation, i.e.,

$$\nabla^2 \phi_{III} = \kappa^2 \phi_{III} \tag{5.40}$$

Kirkwood has shown that the general solution to this equation in spherical coordinates can be written as[44],

$$\phi_{III}(r,\theta,\psi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} \left(\frac{A_{nm}}{r^{n+1}}\right) e^{-\kappa r} K_n(\kappa r) P_n^m(\cos\theta) e^{im\psi} \tag{5.41}$$

[44]Kirkwood, J. G. *J. Chem. Phys.* **1934**, 7, 351.

where,

$$K_n(\kappa r) = \sum_{s=0}^{n} \frac{2^s n!\,(2n-s)!}{s!\,(2n)!\,(n-s)!}\,(\kappa r)^s \qquad (5.42)$$

In order to determine the specific solution for the electrostatic potential given a particular distribution of point charges within the protein it is necessary to implement boundary conditions on equations 5.38, 5.39, and 5.41. The boundary conditions can again be written as equations 5.20 and 5.21 recognizing that in this case there is θ and ψ dependence of the potential. The result of implementing these boundary conditions in the case of a protein charge distribution gives an electrostatic potertial in region I which can be written as[45],

$$\phi(\boldsymbol{x}_1) = \frac{q}{R}\sum_k z_k(A_{k1} - B_{k1}) - \frac{q}{a}\sum_k z_k C_{k1} \qquad (5.43)$$

where,

$$A_{k1} = \frac{R}{\varepsilon_{prot} r_{k1}} \qquad (5.44)$$

$$B_{k1} = \frac{1 - 2\delta}{\varepsilon_{prot}(1-2\rho_{k1}\cos\theta_{k1}+\rho_{k1}^2)^{.5}}$$

$$+ \frac{1}{\varepsilon_{solv}\rho_{k1}} \ln\left(\frac{(1-2\rho_{k1}\cos\theta_{k1}+\rho_{k1}^2)^{.5}+\rho_{k1}-\cos\theta_{k1}}{1-\cos\theta_{k1}}\right) \qquad (5.45)$$

---

[45]Tanford, C.; Kirkwood, J. G. *J. Am. Chem. Soc.* **1957**, 79, 20.

having defined,

$$\delta = \frac{\varepsilon_{prot}}{\varepsilon_{solv}} \tag{5.46a}$$

and

$$\rho_{kl} = \frac{r_k r_l}{R^2} \tag{5.46b}$$

The C_{kl} terms are given by,

$$C_{kl} = \frac{1}{\varepsilon_{solv}} \left\{ \frac{\kappa a}{1+\kappa a} + \sum_{n=0}^{\infty} \frac{2n+1}{2n-1} \left[\frac{\varepsilon_{solv}}{(n+1)\varepsilon_{solv}+n\varepsilon_{prot}} \right]^2 x \right.$$

$$\left. \frac{\kappa a^2 \ \sigma_{kl}{}^n \ P_n(\cos(\theta_{kl}))}{\frac{K_{n+1}(\kappa a)}{K_{n-1}(\kappa a)} + \frac{n(\varepsilon_{solv}-\varepsilon_{prot})}{(n+1)\varepsilon_{solv}+n\varepsilon_{prot}} \left(\frac{R}{a}\right)^{2n+1} \frac{(\kappa a)^2}{4n^2-1}} \right\} \tag{5.47}$$

with,

$$\sigma_{kl} = \frac{r_k r_l}{a^2} \tag{5.48}$$

where, z_i refers to the charge at site i in units of fundamental charge, q is the fundamental charge, r_i is the distance to the charge at site i from the center, r_{kl} is the distance between charges at the kth and lth sites, R is the radius of the protein, a is the ion exclusion radius, θ_{kl} is the angle between the vector position of the charges at sites k and l, κ is the inverse screening length, and ε is the

dielectric constant of the protein or solvent as per subscript.

The A_{kl} terms represent the contribution to the electrostatic potential resulting directly from the bound point charges within the protein. The terms involving B_{kl} represents the correction arising from the bound surface charge at the dielectric boundary. The C_{kl} terms involving the contribution to the potential which arises from the free charge associated with the ions in solution. As expected, this term goes to zero as ionic strength goes to zero.[46]

The work of charging a protein in the Kirkwood-Tanford model can be readily determined at this point. In general, the work of assembling a charge distribution can be written as[47],

$$W_{el} = \frac{q}{2} \sum_{k} z_k \phi(\boldsymbol{x}_k) \tag{5.49}$$

where z_k is the charge at position $\boldsymbol{x}_k$ in terms units of fundamental charge, q. By making use of the electrostatic potential as given by equation 5.42, this can be rewritten as,

$$W_{el} = \frac{q^2}{2R} \sum_{k} \sum_{l \neq k} z_k z_l (A_{kl} - B_{kl}) - \frac{q^2}{2a} \sum_{k} \sum_{l} z_k z_l C_{kl} \tag{5.50}$$

---

[46]Tanford, C.; Kirkwood, J. G. *J. Am. Chem. Soc.* **1957**, 79, 20.
[47]Kirkwood, J. G. *J. Chem. Phys.* **1934**, 7, 351.

The A_{kl} and B_{kl} terms with $k=l$ are self energy terms of the individual charges. Accordingly these terms contain singularities which need to be isolated and recognized as independent of the charge distribution. These terms have been removed from equation 5.50. The C_{kk} terms represent the excess chemical potential of a specific charge resulting from interactions with the ions in solution while C_{kl} terms represent corrections to pairwise interactions resulting from ions in solution.

5.2.3.1 Implementing the Tanford-Kirkwood Model to Study γ_{II}

Implementing the Tanford-Kirkwood model in order to study the titration behavior of a protein requires several additional steps. First, the question of where to place charges on the basis of the x-ray crystallographic structural data must be addressed. In practice, it is not necessary to represent the location of each chargeable residue within a sphere. Instead, a single pair interaction can be considered as a function of separation.[48] From this calculated pair interaction, as a function of distance, all of the charge interactions can be constructed using crystallographically

[48]Tanford, C.; Roxby, R. *Biochem.* **1972**, 11, 11, 2192.

determined distances between chargeable residues.[49,50,51]

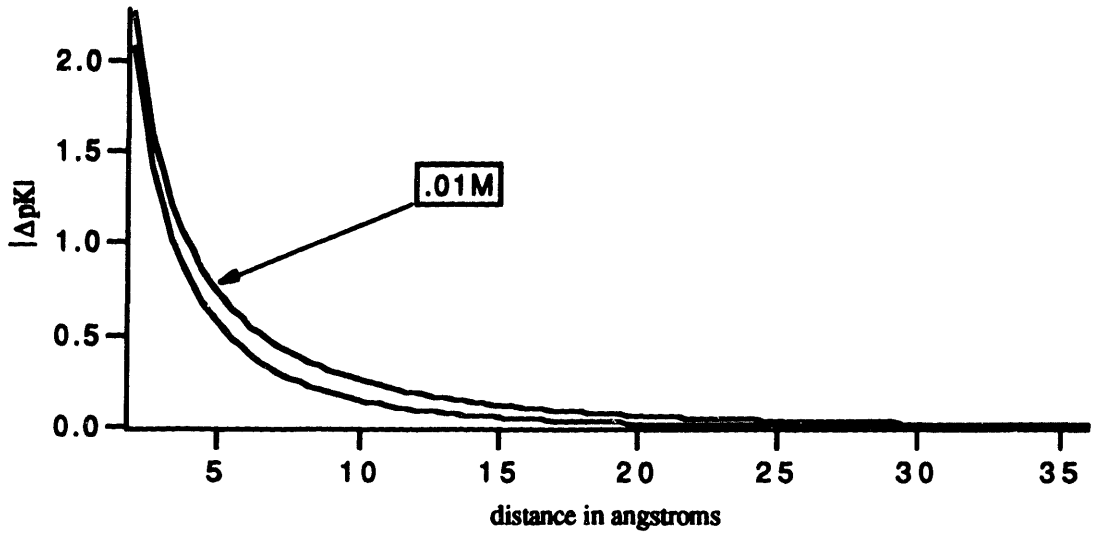


Figure 5.18 -- The pair interaction energy in terms of ΔpK as a function of separation distance in the Tanford-Kirkwood model. The charge is assumed to be at the dielectric interface. The protein radius is 18Å and the ion exclusion radius is 20 Å. The upper curve is for ionic strength of 0.01 M and the lower curve 0.1 M.

The calculated interaction energy for a single charge pair given protein dimensions appropriate to γ_{II} and 0.1 M KCl solvent is presented in 5.13, adapted from Matthew et al.[52] In this figure the depth of the charges is taken to be zero the protein radius 18 Å, and the ion exclusion radius 20 Å.

---

[49]Imoto, T. *Biophys. J.* **1983**, 44, 193.
[50]Friend, S.; Gurd, F. *Biochem.* **1979**, 21, 4612.
[51]Matthew, J. B.; Richards, F. M. *Biochem.* **1982**, 21, 4989.
[52]Matthew, J. B.; Hanania, G.; Gurd, F. *Biochem.* **1979**, 18, 10, 1919.

5.2.3.2 Solvent Accessibility and the Modified Tanford-Kirkwood Model

Tanford and Roxby[53] were the first to recognize that amino acid residues at the protein surface often protrude into the solvent, and hence do not feel the effects of the internal dielectric constant as fully as suggested by the Kirkwood-Tanford model. In order to address this issue, Shire, Hanania, and Gurd introduced a method whereby the accessibility of each group to solvent is taken into account.[54] In this approach, the dielectric constant which is appropriate for charge interactions within the protein is recognized to be dependent on the extent to which an interacting residue is exposed to solvent. In other words, for residues which are fully exposed to solvent, the appropriate dielectric constant will be closer to that of water than that of the protein interior. In order to implement this, a simple correction to the charge pair interaction at zero depth is imposed. The new charge pair interaction is taken to be,

$$W'_{ij} = W_{ij}(r_{ij}) \left(1 - \frac{SA_i+SA_j}{2}\right) \qquad (5.51)$$

where $W_{ij}(r_{ij})$ is the interaction energy for charges separated

---

[53]Tanford, C.; Roxby, R. *Biochem.* **1972**, 11, 11, 2192.
[54]Shire, S. J.; Hanania, G.; Gurd, F. *Biochem.* **1974**, 13, 2967.

by a distance r_{ij} at the protein surface (see figure 5.13). *SA* is the solvent accessibility parameter. It is defined as the area of an amino acid residue which is exposed to solvent in the protein normalized to the area of the same residue in a short, straight peptide (Ala-X-Ala). Lee and Richards studied and catalogued the area of titrateable amino acid residues exposed to solvent in short peptides in 1971.[55] The area of a residue which is available to solvent can be determined from x-ray crystallographic studies. In the case of γ_{II}, these areas have been determined by Wistow.[56] Introduction of this modification to the Kirkwood-Tanford model adequately accounts for variable charge depth.[57,58,59,60]

Equation 5.51 is not an exact expression. Since the exact dependence of a charge pair interaction on solvent accessiblities is very difficult to determine, the simplest, linear model is employed. The choice of this linear form is justified by recognizing that the electrostatic interaction energy varies inversely with the dielectric constant. The solvent accessibilities are indirect measures of the effective local dielectric constant. As the solvent accessibility varies from zero to one, the effective dielectric constant varies from that of the protein interior

---

[55]Lee, B.; Richards, F. M. *J. Mol. Biol.* **1971**, 55, 379.
[56]Wistow, G. *J. Mol. Bio.* **1983**, 170, 175.
[57]Friend, S.; Gurd, F. *Biochem.* **1979**, 21, 4612.
[58]Glackin, M.; et. al. *Prot.* **1989**, 5, 66.
[59]Imoto, T. *Biophys. J.* **1983**, 44, 193.
[60]Matthew, J. B.; Richards, F. M. *Biochem.* **1982**, 21, 4989.

to that of the solvent, which is much larger. In the method of Shire, Hanania, and Gurd the effective dielectric constant at a residue is assumed to vary inversely with the solvent accessibility. Furthermore, the effective dielectric constant for interaction between two sites is taken to be the average of the effective dielectric constants.

5.2.3.3 Electrostatic Corrections to the *pKs* of γ_{II} Using the Modified Tanford-Kirkwood Model

Software to implement the Tanford-Kirkwood model for the case of γ_{II} was written in C and executed on a DEC Micro VAX II. It is reproduced in appendix F. The Brookhaven Protein Data Bank file on γ_{II} was evaluated and the distances between all of the titrateable residues was calculated. For residues where charge could be distributed over more than one atom, i.e. carboxylic acid, the location was taken to be the average of the positions of the potentially charged sites. The solvent accessibility parameter, *SA*, of each site was then computed. The area of each residue exposed to solvent was taken from Wistow[61], and the reference area taken from Lee and Richards[62] as described in section 5.2.3.2. The resulting solvent accessibility parameters for γ_{II} are presented in

---

[61]Wistow, G. *J. Mol. Bio.* **1983**, 170, 175.
[62]Lee, B.; Richards, F. M. *J. Mol. Biol.* **1971**, 55, 379.

Table 5.1. The *pK* of each residue was then modified by summing over all pair interactions, corrected for solvent accessibility[63,64], i.e.,

$$\Delta pK_i = - (kT \ln 10)^{-1} \sum_{j \neq i} W'_{ij} \; z_j \qquad (5.52)$$

where z_j is the fractional charge at site j , W'_{ij} is the pair interaction energy of two unit charges including corrections for solvent accessibility, k is Boltzmann's constant, and T is temperature.

Since each pair interaction was weighted by the charges at the pair sites, the electrostatic corrections were evaluated in an iterative fashion at each of 100 *pH* values between 2 and 12. In the first iteration at each *pH* the charges on the residues were calculated using the uncorrected, intrinsic *pK*s as given in Table 5.2. In subsequent iterations the *pK*s used to determine the charge states of the residues were corrected by the prior iteration. Iterations were repeated at each *pH* until the change in corrected *pK*s from one iteration to the next was insignificant. The routine was rapidly convergent, with less than 10 iterations required at each *pH*. The *pK*s predicted from this process for γ_{II} are presented in Table

---

[63]Garcia-Moreno, B.; et. al. *J. Biol. Chem.* **1985**, 260, 26, 14070.
[64]Shire, S. J.; Hanania, G.; Gurd, F. *Biochem.* **1974**, 13, 2967.

5.2.

Table 5.1*

Residue	pK_a	Position (Å)			SA
		x	y	z	
1 GLY:	10.4	-2.14	6.665	31.887	.51
2 LYS:	10.4	-3.06	27.82	40.797	.52
7 GLU:	4.5	1.00	23.27	47.59	.11
8 ASP:	4.0	5.96	21.5	50.94	.36
9 ARG:	12	15.8	21.7	47.1	.73
14 HIS:	6.4	-2.91	20.1	47.0	.60
15 CYS:	8.6	-.75	14.1	47.0	.95
16 TYR:	10.0	-6.57	17.8	38.3	.31
17 GLU:	4.5	-3.83	9.54	42.5	.68
21 ASP:	4.0	-2.44	15.8	26.4	.61
28 TYR:	10.0	-8.82	20.7	37.4	.57
31 ARG:	12	4.03	24.8	49.6	.43
36 ARG:	12	6.96	8.6	47.0	.41
38 ASP:	4.0	2.79	5.5	40.95	.70
41 CYS:	8.6	6.91	7.37	27.8	.19
46 GLU:	4.5	9.85	28.9	32.91	.26
47 ARG:	12	4.6	32.3	27.68	.48
53 HIS:	6.4	13.5	27.5	31.28	.26
58 ARG:	12	13.6	6.03	28.28	.26
59 ARG:	12	6.92	2.04	39.13	.20
61 ASP:	4.0	9.31	6.83	43.75	.57
64 ASP:	4	16.5	18.1	43.95	.27
65 TYR:	10	15.4	24.9	42.2	.32
73 ASP:	4	9.14	26.6	43.59	.37
76 ARG:	12	3.31	29.5	39.27	.40
79 ARG:	12	.97	20.3	23.33	.27
84 HIS:	6.4	4.20	5.17	15.68	.39
89 ARG:	12	10.62	.22	7.64	.39
91 ARG:	12	9.12	12.2	3.4	.37
93 TYR:	10	11.5	17.3	3.91	.30
94 GLU:	4.5	22.4	22.6	3.13	.58
95 ARG:	12	20.2	25.2	.32	.48
96 ASP:	4	13.7	29.7	6.44	.47
97 ASP:	4	8.67	27.23	3.39	.57
99 ARG:	12	15.86	29.11	.98	.47
104 GLU:	4.5	12.31	8.76	5.86	.45
107 ASP:	4	19.17	5.35	14.45	.58
108 ASP:	4	15.81	6.7	20.9	.39

*Tyrosines and cysteines with solvent accessibilities below 15% were considered to be buried within the protein and hence not to titrate. It is commonly observed that these types of residues do not participate in acid-base titration when buried.

Table 5.1(continued)

Residue	pK_o	Position (Å)			SA
114 ASP:	4	30.21	13.10	17.544	.81
115 ARG:	12	23.77	5.78	11.51	.79
117 HIS:	6.4	30.0	14.3	9.35	.95
120 GLU:	4.5	22.67	27.76	13.51	.95
122 HIS:	6.4	18.66	27.66	10.93	.33
128 GLU:	4.5	6.24	6.14	9.73	.47
135 GLU:	4.5	20.47	25.23	25.1	.22
139 TYR:	10	20.1	9.25	22.1	.28
140 ARG:	12	26.91	18.22	26.87	.60
142 ARG:	12	16.73	29.24	27.95	.39
147 ARG:	12	.77	19.32	20.38	.46
150 GLU:	4.5	4.25	20.51	5.91	.51
152 ARG:	12	6.44	23.2	3.73	.56
153 ARG:	12	7.08	31.2	12.93	.59
154 TYR:	10	15.3	32.1	16.1	.25
156 ASP:	4	6.68	26.87	12.79	.50
163 LYS:	10.4	24.93	26.98	17.63	.33
168 ARG:	12	17.02	10.17	26.55	.29
169 ARG:	12	14.36	3.46	18.42	.25
172 ASP:	4	6.83	3.37	29.8	.42
174 TYR:	10	3.55	-2.03	25.6	.95
174 TYR:	3.6	9.16	-3.15	29.44	.95

In Table 5.2, the calculated *pK*s are boldfaced when they are being calculated for *pH* close to their *pK* value. The corrected *pK*s are functions of *pH*. The transition from binding to nonbinding, however, occurs when the *pH* is near the *pK*. Accordingly, the calculated *pK* for the *pH* nearest the *pK* is the effective *pK*.

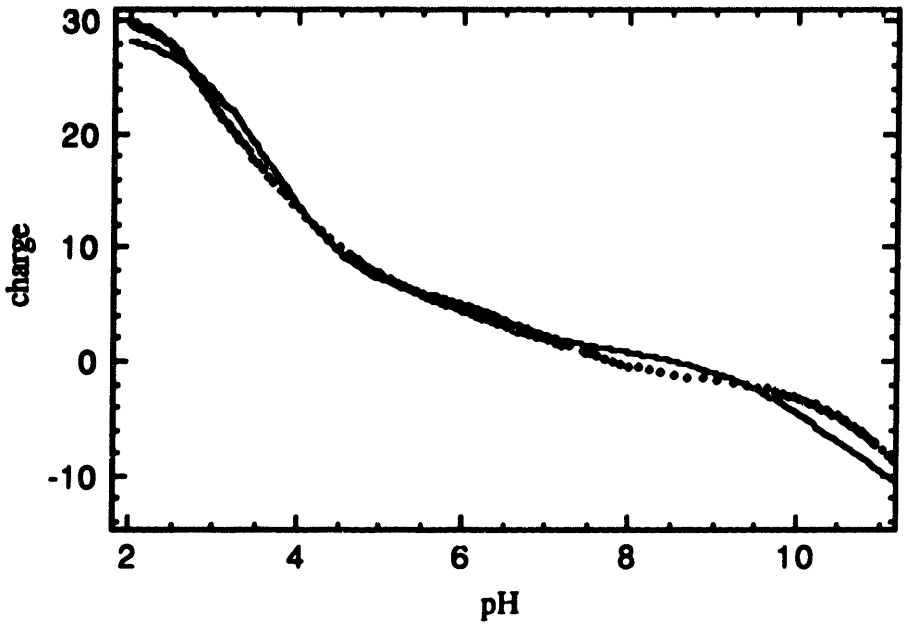


Figure 5.19 -- Theoretical titration curve of γ_{II} as predicted by the modified Tanford-Kirkwood model. Experimental data is plotted as dots for comparison.

The theoretical titration curve resulting from these *pK*s is shown in figure 5.19 with titration data for comparison. As figure 5.19 demonstrates, the theoretical titration curve is in very good agreement with the experimentally determined titration curve. The close agreement between the predicted and observed titration curves lends support to the legitimacy of the predicted *pK*s in table 5.2.

Table 5.2

Residue	pKo	Effective pK	pH=2	pH=4	pH=6	pH=8	pH=10
1 GLY	10	10.7	10.2	10.4	10.4	10.5	10.7
2 LYS	10	11.2	10.3	10.9	11	11.1	11.2
7 GLU	4.5	3.82	3.43	3.82	4.08	4.28	4.43
8 ASP	4	3.74	3.37	3.74	3.83	3.9	3.99
9 ARG	12	12.6	12	12.3	12.4	12.4	12.6
14 HIS	6.4	6.7	6.3	6.69	6.73	6.8	6.95
15 CYS	8.6	8.7	8.39	8.57	8.64	8.68	8.76
16 TYR	10	10.5	9.7	9.91	9.99	10.1	10.5
17 GLU	4.5	3.86	3.72	3.86	3.9	3.98	4.31
21 ASP	4	3.6	3.51	3.6	3.64	3.69	3.87
28 TYR	10	10.1	9.83	9.94	9.99	10.1	10.4
31 ARG	12	13.2	12.1	12.9	13.1	13.2	13.2
36 ARG	12	12.6	11.9	12.4	12.4	12.5	12.6
38 ASP	4	3.64	3.43	3.64	3.67	3.72	3.87
41 CYS	8.6	8.7	7.97	8.56	8.66	8.7	8.91
46 GLU	4.5	3.63	3.43	3.63	4.04	4.2	4.35
47 ARG	12	12.1	11.7	11.9	12	12	12.1
53 HIS	6.4	6.52	5.86	6.46	6.52	6.55	6.71
58 ARG	12	12.4	11.5	11.9	11.9	12.1	12.4
59 ARG	12	12.6	11.8	12.3	12.4	12.4	12.6
61 ASP	4	3.56	3.36	3.56	3.59	3.63	3.71
64 ASP	4	3.73	3.52	3.73	3.78	3.82	4.02
65 TYR	10	10.3	9.63	10.1	10.2	10.2	10.3
73 ASP	4	3.84	3.54	3.84	3.93	3.98	4.21
76 ARG	12	12.4	11.8	12.1	12.2	12.3	12.4
79 ARG	12	11.7	11.1	11.4	11.5	11.5	11.7
84 HIS	6.4	6.48	6.11	6.44	6.48	6.53	6.64
89 ARG	12	12.6	11.7	12.3	12.5	12.5	12.6
91 ARG	12	12.6	11.8	12.2	12.3	12.4	12.6
93 TYR	10	9.94	9.34	9.77	9.87	9.89	9.94
94 GLU	4.5	4.05	3.93	4.05	4.13	4.17	4.27
95 ARG	12	12.3	11.7	12.1	12.2	12.2	12.3
96 ASP	4	3.64	3.34	3.64	3.8	3.88	4.04
97 ASP	4	3.76	3.46	3.76	3.82	3.85	3.97
99 ARG	12	12.3	11.7	12	12.1	12.2	12.3
104 GLU	4.5	3.96	3.72	3.96	4.03	4.06	4.21
107 ASP	4	3.68	3.49	3.68	3.73	3.77	3.91
108 ASP	4	3.21	2.94	3.21	3.28	3.39	3.77
114 ASP	4	3.9	3.84	3.9	3.93	3.95	4.02
115 ARG	12	12.2	11.9	12.1	12.1	12.1	12.2
117 HIS	6.4	6.41	6.32	6.39	6.41	6.41	6.45
120 GLU	4.5	4.09	3.98	4.09	4.25	4.33	4.54
122 HIS	6.4	6.65	6.05	6.57	6.65	6.68	6.95
128 GLU	4.5	3.99	3.77	3.99	4.17	4.24	4.33
135 GLU	4.5	3.87	3.7	3.87	4.01	4.08	4.29
139 TYR	10	9.99	9.37	9.79	9.85	9.9	9.99
140 ARG	12	12.1	11.8	12	12	12	12.1

142 ARG	12	**12.4**	11.4	11.9	12.2	12.3	12.4
147 ARG	12	**11.7**	11.1	11.4	11.4	11.5	11.7
150 GLU	4.5	**4.15**	3.92	4.15	4.2	4.23	4.38
152 ARG	12	**12.8**	11.9	12.5	12.6	12.6	12.8
153 ARG	12	**12.5**	11.9	12.3	12.3	12.4	12.5
154 TYR	10	**10.1**	9.44	9.76	9.94	10	10.1
156 ASP	4	**3.61**	3.36	3.61	3.68	3.72	3.87
163 LYS	10	**10.8**	10.1	10.5	10.7	10.7	10.8
168 ARG	12	**12.4**	11.5	11.8	11.9	12	12.4
169 ARG	12	**12.8**	11.8	12.4	12.5	12.6	12.8
172 ASP	4	**3.67**	3.46	3.67	3.73	4.04	4.39
174 TYR	10	**10.1**	9.86	9.97	10	10.1	10.1
174 TYR	3.6	**3.55**	3.44	3.55	3.57	3.61	3.67

Chapter 6 - Summary and Future Work

6.1 Summary and Conclusions

This work represents the first thorough exploration of the charge distribution on an important family of phase-separating ocular lens proteins, the γ-crystallins. The specific questions addressed and answered were: (1) What is the net charge on these proteins as a function of pH?; (2) Do ionic strength and electrolyte identity effect the net protein charge?; (3) Can *pK*s and titration curves which agree with the experimental results be theoretically predicted?; (4) What is the work associated with charging these proteins, and finally; (5) What is the electrostatic interaction energy between the proteins and solvent.

In order to determine the net charge as a function of *pH* on these proteins, acid-base titration experiments were performed. The titration apparatus which was developed is fully automated and utilizes a combined glass membrane electrode. Two methods were developed to convert the raw acid-base titration data to protein charge as a function of

pH. The accuracy of these techniques was demonstrated over a wide *pH* range (2 to 12) by performing titration experiments on simple molecules whose proton binding properties are known.

The titration curves of the γ-crystallins were studied most extensively in 0.1M KCl. They were remarkably similar for all four of the protein fractions studied with the exception of the basic range titration of γ_{IIIb}. Above *pH* 9 the titration curve of γ_{IIIb} fell below that of the other proteins. This difference is most pronounced slightly above *pH* 10 where there are four protons fewer bound to γ_{IIIb}, and hence the net charge of γ_{IIIb} is four larger in magnitude. It is also interesting that γ_{IIIb} is the only fraction with a different isoelectric point. Its isoelectric point is 7.4 while the other fractions have a common isoelectric point of 7.8.[65]

The effect on the titration curves of varying the ionic conditions of the solvent was studied. It was found that varying the identity of the electrolyte in solution had no effect on the titration curves of either γ_{IIIb} or γ_{IV}. Both cation and anion were varied, however, the electrolytes which were examined were all monovalent, i.e. KCl, NaCl, NaBr. An effect on the titration curves which resulted from varying the ionic strength of the solvent was observed. Titration experiments were performed with γ_{IV} in both 0.1M KCl and 0.01M

---

[65]McDermott, M. J.; et. al. *Arch. Biochem. Biophys.* **1988**, 262, 609.

KCl. It was found that reducing the ionic strength resulted in a slight flattening of the titration curves which corresponds to a reduction in net charge at a given *pH*. This result is consistent with expectations. By reducing the ionic strength the screening of the electric field from the protein is reduced. Reducing ionic strength makes a given net protein charge less energetically favorable, resulting in a shift of *pK*s in order to reduce the magnitude of the charge at fixed *pH*. This effect would be expected to become more pronounced at *pH*s where the net protein charge was larger. Indeed, the titration curves in 0.1M KCl and 0.01M KCl are essentially identical between *pH* 4.5 and 9 and become progressively more divergent as the *pH* extremes are approached.

In order to understand the titration curves theoretically it was necessary to correct the intrinsic *pK*s of the protein for electrostatic interactions that alter the proton binding energies. To correct *pK*s for electrostatic interactions, the electrostatic potential at each titrateable residue was required. The potential was determined using three methods. The first method, attributable to Linderstrom-Lang, treats the protein as a uniform spherical charge shell and linearizes the Poisson-Boltzmann equation in describing the electrolyte. The theoretical titration curves which this method produced were in fair agreement with experiment over a *pH* range restricted to the vicinity of the

isoelectric points. It demonstrated overestimation of the *pK* corrections in the *pH* extremes, resulting in an overly flattened titration curves. The performance of the Linderstrom-Lang method in predicting the titration curves was comparable at ionic strength 0.1M and 0.01M. There are two significant shortcomings of the Linderstrom-Lang method. First, the linearization of the Poisson-Boltzmann equation is not legitimate at the protein surface except for low net charge, i.e. near the isoelectric point. Second, *pK* corrections can not be assigned to specific residues.

The second method which was used to theoretically predict the titration curves differed from the Linderstrom-Lang method by utilizing the nonlinearized Poisson-Boltzmann to describe the electrolytic solvent. This method produced theoretical titration curves which were in very good agreement with the data obtained for γ_{II} and γ_{IV}. This method, despite its accuracy in reproducing the observed titration curves, does not allow for determining the location of charge within the protein.

The third method used to predict protein titration curves was that of Kirkwood and Tanford. In this method, charge at titrateable sites is treated as a point charge and assigned to specific locations within the protein, but the electrolyte is still described with the linearized Poisson-Boltzmann equation. This method not only predicts the titration curve, but also predicts the *pK* of each titrateable

residue within the protein. To do this, refined structural data is required for the protein. In addition to the spatial location of each titrateable residue within the protein, information about the accessibility to the protein surface of each titrateable residue is also needed. Adequate information to fully implement the Kirkwood-Tanford was only available for γ_{II}. This method produced a theoretical titration curve that was in excellent agreement with experimental results for γ_{II}. It also produced predictions regarding the exact location of charge as a function of *pH* for γ_{II}.

Both the nonlinearized modification to the Linderstrom-Lang model and the Kirkwood-Tanford model predict the experimental titration curves equally well. This is of particular interest given the very different assumptions involved in these two approaches. In the nonlinearized Linderstrom-Lang model the protein charge is treated as though it were a spherical shell. The details of the location of the charge on the surface of the protein do not enter into the model. The electrolyte is treated with the accurate, nonlinearized Poisson-Boltzmann equation. In the Kirkwood-Tanford model the location of charge within the protein is treated in detail, however the electrolyte is described with the linearized, and hence less accurate, Poisson-Boltzmann equation.

The success of the nonlinearized Linderstrom-Lang model suggests that the location of charge within the protein, as determined crystallographically, may not be important in determining the proton binding properties of the protein. The effect of time averaging conformational fluctuations of the protein may lessen the importance of the location of charge. Fluctuations in the instantaneous position of charge may be sufficiently large so that the charge is effectively smeared over the protein surface.

In contrast to the nonlinearized Linderstrom-Lang model, the success of the Kirkwood-Tanford model suggests that the location of protein charge, as determined crystallographically, significantly effects proton binding. The Kirkwood-Tanford model successfully reproduces the experimental results despite its crude treatment of the electrolyte in which the linearized Poisson-Boltzmann equation is used.

It is not feasible to solve a nonlinearized version of the Kirkwood-Tanford model. The significant changes which resulted from using the nonlinearized version of the Linderstrom-Lang model suggest that were it feasible to solve the nonlinear Kirkwood-Tanford model, the results may differ considerably from those of the linearized version. Because the agreement of the linearized Kirkwood-Tanford model with experiment is so good, the possibility that improving the description of the electrolyte may worsen the fit. It is

conceivable that the errors introduced by using the linearized Poisson-Boltzmann equation may cancel errors introduced by over specifying the location of the charge.

Calculations of the electrostatic interaction energy between protein and solvent were done assuming a uniform spherical shell charge distribution (the Linderstrom-Lang model). Using this method it was found that the change in electrostatic solvation energy of a lysozyme protein molecule, which resulted from changing the charge, was much larger than that expected on the basis of charge versus cloud point measurements performed by Taratuta.[66] This suggests that the cloud point temperature can not be understood simply in terms of the protein solvation energy. Because the protein-solvent interaction energy primarily involves solvent molecules close to the protein surface, it probably depends weakly on protein concentration. Other charge dependent contributions to the cloud point, such as protein-protein interactions, may play a more important role and must be included in a successful theory of protein liquid-liquid phase separation.

---

[66]Taratuta, V.; et. al. *J. Phys. Chem.* **1990**, 5, 2140.

6.2 Future Work

A long term objective of this work has always been to understand theoretically the effect of protein charge on the critical temperature of protein solutions. In order to achieve this goal it was necessary to fully characterize these proteins' charge distribution as it depends on solution conditions. That has now been done. The next step is a careful experimental determination of how the critical temperature depends on solution conditions such as *pH*. Some preliminary experiments addressing this question in the protein Lysozyme have been performed by Taratuta.[67] Lysozyme's charge distribution has been studied in considerable detail as well.[68] With the completion of this work, there is now detailed information regarding the charge distribution in several phase separating proteins. This information, together with a thorough experimental investigation of the dependence of critical temperature on solvent conditions should prove adequate to guide theoretical attacks on this interesting problem.

Another related extension of this work has to do with ongoing efforts to alter the critical temperature of protein solutions by protein modification. The question as to whether or not the observed changes in T_c are due to changes

---

[67]Taratuta, V.; et. al. *J. Phys. Chem.* **1990**, 5, 2140.
[68]Imoto, T. *Biophys. J.* **1983**, 44, 193.

in protein charge often seems to arise. While the resolution to this question is rooted in a deeper understanding of the effects of protein charge on T_c, it is necessary to know what the charge is on these modified proteins. With the apparatus and analysis techniques developed in this work, it should be relatively easy and quick to determine the charge of the modified proteins as it depends on solution conditions.

Appendix A - Dispenser Control Software

This software was written in C by Michael Orkiscz. Its functions include monitoring pH and activating the acid/base dispenser.

```
#include <stdio.h>
#include "data.h"
#include "odds.h"

#define MIN_VOLUME  100

struct devices
{
      FILE *output;
      int dispenser;
      int meter;
};

static void
      check_and_display_settings(), display_increments(),
      test_and_initialize_devices(), titrate(), dispense();
static double
      get_pH();

int keyboard;            /* Channel for reading the keyboard
*/
                    /* (this shouldn't be needed explicitly)
*/

main(argc, argv)
   int argc;
   char *argv[];
{
      char fname[256], *strcpy();
      struct data data;
      struct devices devices;
      FILE* script;
      void parse();

      /*
       *  First get the name and open the script file.
       *  If it is not the first command-line argument,
       *  ask for it explicitly.  Similarly the log file.
       */
```

```
        if (argc >= 2)
           (void) strcpy(fname, argv[1]);
        else
        {
           type("script file: ");
           (void) scanf(" %s", fname);
        }
        if (!(script = fopen(fname, "r")))
           error("Cannot open %s\n", fname);

        if (argc >= 3)
           (void) strcpy(fname, argv[2]);
        else
        {
           type("log file [NONE]: ", fname);
           (void) fflush(stdin);
           scanf("%*[ \t]%[^ \t\n]", fname);
        }
        if (*fname)
           if (!(logfile = fopen(fname, "w")))
              error("Cannot open %s\n", fname);

        if ((keyboard = open_term("sys$input")) == -1)
           error("Couldn't open standard input!!! -
HELP!!!\n");

        /*
         *  Initialize the data structure.
         *  Everything is unassigned.
         */

        data.auto_stop.status = UNASSIGNED;
        data.dispenser = NULL;
        data.increments.array = NULL;
        data.initialization = NULL;
        data.output = NULL;
        data.ph_meter = NULL;
        data.ph_samples = UNASSIGNED;
        data.probing_time = UNASSIGNED;
        data.refill_point = UNASSIGNED;
/*      data.sampling_time = UNASSIGNED;  */
        data.slowdown = UNASSIGNED;
        data.speed = UNASSIGNED;
        data.time_limit = UNASSIGNED;
        data.tolerance = UNASSIGNED;
        data.volume = UNASSIGNED;

        /*  Get the data from the script file    */

        parse(script, &data);

        /*  Make sure everything is assigned and in range
*/
```

```
        check_and_display_settings(&data, &devices);

        /*  Make sure devices are responding properly      */

        test_and_initialize_devices(&data, &devices);

        /*  Gentelmen, start your engines ... Go!          */

        titrate(&data, &devices);

        message(MSG_NOWAIT, "\tCONGRATULATIONS  --  we are
done!\n");
}


static void
check_and_display_settings(data_ptr, dev_ptr)
    struct data *data_ptr;
    struct devices *dev_ptr;
{
        type("Data assigned as follows:\n\n");

        if (data_ptr->output == NULL)
            type("\tNo output file specified.\n");
        else {
            type("\tOutput directed to file %s\n", data_ptr-
>output);
            if (!(dev_ptr->output = fopen(data_ptr->output,
"w")))
                error("Cannot open %s for writing\n", data_ptr-
>output);
        }

        if (data_ptr->dispenser == NULL)
            type("\tDispenser assigned default device %s\n",
                data_ptr->dispenser = DFLT_DISPENSER);
        else type("\tDispenser device = %s\n", data_ptr-
>dispenser);
        if ((dev_ptr->dispenser = open_term(data_ptr-
>dispenser)) == -1)
            error("Cannot open device %s\n", dev_ptr-
>dispenser);

        if (data_ptr->ph_meter == NULL)
            type("\tpH-meter assigned default device %s\n",
                data_ptr->ph_meter = DFLT_PH_METER);
        else type("\tpH-meter device = %s\n", data_ptr-
>ph_meter);
        if ((dev_ptr->meter = open_term(data_ptr->ph_meter)) ==
-1)
            error("Cannot open device %s\n", dev_ptr->meter);

        type("\n");
```

```
      if (data_ptr->initialization == NULL)
         type("\tNo explicit initialization\n");
      else
      {
         upcase(data_ptr->initialization);
         if (!strcmp(data_ptr->initialization, PRELOAD))
         {
            type("\tInitialization using default preload
sequence\n");
            data_ptr->initialization = DFLT_INITIALIZATION;
         }
         else type("\tInitialization = \"%s\"\n",
               data_ptr->initialization);
      }

      if (data_ptr->speed == UNASSIGNED)
         type("\tPlunger speed unaffected\n");
      else if (data_ptr->speed == 0)
         type("\tPlunger speed set for external
regulation\n");
      else if (data_ptr->speed <= 15)
         type("\tPlunger speed = %d\n", data_ptr->speed);
      else error("Plunger SPEED must be 1-15 or 0 for
external regulation\n");

      if (data_ptr->slowdown == UNASSIGNED)
         type("\tPlunger slowdown unaffected\n");
      else if (data_ptr->slowdown <= 99)
         type("\tPlunger slowdown = %d steps\n", data_ptr-
>slowdown);
      else error("Plunger SLOWDOWN must be 0-99\n");
      type("\n");

      if (data_ptr->probing_time == UNASSIGNED)
         type("\tProbing time assigned default value %d s\n",
             data_ptr->probing_time = DFLT_PROBING_TIME);
      else type("\tProbing pH every %d s\n", data_ptr-
>probing_time);
      if (data_ptr->probing_time > 0) --data_ptr-
>probing_time;

      if (data_ptr->ph_samples == UNASSIGNED)
         type("\tTaking default %d pH samples per data
point\n",
             data_ptr->ph_samples = DFLT_PH_SAMPLES);
      else type("\tTaking %d pH samples per data point\n",
            data_ptr->ph_samples);

      if (data_ptr->tolerance == UNASSIGNED)
         type("\tTolerance assigned default value %.3f pH\n",
             data_ptr->tolerance = DFLT_TOLERANCE);
      else type("\tTolerance = %.4f pH\n", data_ptr-
>tolerance);
```

```
      if (data_ptr->time_limit == UNASSIGNED)
         type("\tNo time limit for pH stabilization\n");
      else type("\tTime for pH stabilization limited to %d
s\n",
            data_ptr->time_limit);

      if (data_ptr->auto_stop.status == UNASSIGNED)
         type("\tNo pH level for AUTO STOP specified\n");
      else if (data_ptr->auto_stop.level < 0 || data_ptr-
>auto_stop.level > 14)
         error("The AUTO STOP pH level must be between 0 and
14\n");
      else type("\tAUTO STOP when pH crosses %f\n",
            data_ptr->auto_stop.level);
/*
      if (data_ptr->sampling_time == UNASSIGNED)
         type("\tDefault time between samples is %d s\n",
             data_ptr->sampling_time = DFLT_SAMPLING_TIME);
      else type("\tSamples taken every %d s\n", data_ptr-
>sampling_time);
*/
      type("\n");

      if (data_ptr->volume == UNASSIGNED)
         error("The syringe VOLUME must be assigned\n");
      else if (data_ptr->volume == 0)
         error("Attempt to assign a ZERO syringe VOLUME\n");
      else type("\tSyringe volume = %g units\n", data_ptr-
>volume);

      if (data_ptr->refill_point == UNASSIGNED)
         type("\tRefill point set to default value of
%g%%\n",
             data_ptr->refill_point = DFLT_REFILL_POINT);
      else if (data_ptr->refill_point > 100.0)
         error("REFILL POINT must be between 0%% and
100%%\n");
      else type("\tSyringe refilled after dispensing %g%% of
its volume\n",
           data_ptr->refill_point);
      data_ptr->refill_point = (int) (data_ptr->refill_point
* 10.0 + 0.5);

      if (data_ptr->increments.array == NULL)
         error("The INCREMENTS must be assigned\n");

      wait_for_user();

      type("\n");
      display_increments(data_ptr);
}
```

```
static void
display_increments(data_ptr)
   struct data *data_ptr;
{
      register float *farray = data_ptr->increments.array;
      register double steps_per_unit_volume = 1000.0 /
data_ptr->volume;
      register int index, i, length = data_ptr-
>increments.length;

      type("\tThe increments are (in %%'s of syringe's
volume):\n");

      for (i = 0, index = 0; index < length; ++index)
      {
         if (!i--)       /* if necessary to start a new line
*/
         {
            type("\n   ");
            i = 8 - 1;   /* 10 items per line */
         }
         farray[index] *= steps_per_unit_volume;
         type("%8.1f%%", farray[index] / 10.0);
      }
      type("\n");
}


static void
test_and_initialize_devices(data_ptr, dev_ptr)
   struct data *data_ptr;
   struct devices *dev_ptr;
{
      double pH_probe();
      void dispenser_initialize(), pH_initialize();

      type("\n\tTesting the pH meter connection  ...  ");
      pH_initialize(dev_ptr->meter);

      type("\n\tInitializing the dispenser  ...  ");
      dispenser_initialize(dev_ptr->dispenser,
                  data_ptr->initialization,
                  data_ptr->speed,
                  data_ptr->slowdown);
      type("\n");
      message(MSG_WAIT,
           "\tAll systems OK.\n%s\n%s\n%s\n",
           "\tThis is your last chance to change anything",
           "\tIf you have any objections, state them NOW,",
           "\tor forever hold your peace!");
}


static void
```

```
titrate(data_ptr, dev_ptr)
   struct data *data_ptr;
   struct devices *dev_ptr;
{
      register float *farray = data_ptr->increments.array;
      int length = data_ptr->increments.length;
      int refill = data_ptr->refill_point;
      int total = 0, contents = 0;
      int index, steps;

      void dispenser_refill(), write_data_point();
      double pH, volume, pH_data_point();
      double fsteps = 0.0;

      type("\n    0 - (0.0%%)\t");

      pH = get_pH(dev_ptr->meter, data_ptr);
      write_data_point(dev_ptr->output, 0.0, pH);

      for (index = 0; index < length; ++index)
      {
         /*
          *    Round the number of steps to the nearest
integer
          *    and add the difference to the next one, so the
          *    round-off error does not accumulate.
          */
         fsteps += farray[index];
         steps = fsteps + 0.5;
         fsteps -= steps;

         type("%5d - ", index + 1);

         dispense(steps,
               &contents,
               (int) data_ptr->refill_point,
               dev_ptr->dispenser);

         volume = (total += steps) * (data_ptr->volume /
1000.0);
         pH = get_pH(dev_ptr->meter, data_ptr);
         write_data_point(dev_ptr->output, volume, pH);
      }
}


static void
dispense(steps, vol_ptr, refill, dispenser)
   int steps, *vol_ptr, refill;
   int dispenser;
{
      void dispenser_dispense(), dispenser_refill();

      while (steps + *vol_ptr > refill)
```

```
      {
         if (steps <= refill || refill - *vol_ptr <
MIN_VOLUME)
         {
            /* do nothing but refill */
         }
         else if (steps + *vol_ptr - refill < MIN_VOLUME)
         {
            type("(%.1f%%)\t", (steps - MIN_VOLUME) / 10.0);
            dispenser_dispense(dispenser, steps -
MIN_VOLUME);
            *vol_ptr += steps - MIN_VOLUME;
            steps = MIN_VOLUME;
         }
         else       /* Large volume */
         {
            type("(%.1f%%)\t", (refill - *vol_ptr) / 10.0);
            dispenser_dispense(dispenser, refill - *vol_ptr);
            steps -= refill - *vol_ptr;
            *vol_ptr = refill;
         }
         message(MSG_NOWAIT,
              "\tRefilling dispenser (%g%%)\n", *vol_ptr /
10.0);
         type("\t");
         dispenser_refill(dispenser, *vol_ptr);
         *vol_ptr = 0;
      }
      type("(%.1f%%)\t", steps / 10.0);
      dispenser_dispense(dispenser, steps);
      *vol_ptr += steps;
}


static
get_pH(pH_ptr, meter, data_ptr)
   float *pH_ptr;
   int meter;
   struct data *data_ptr;
{
      register int index, number = data_ptr->ph_samples;
      double deviation(), pH_probe(), pH_data_point();
      int start_time = get_time();
      double sum, stddev;
      float pH[128];
      int i = 5;
      int key;

      for (index = 0; index < number; ++index)
      {
         if (index)
         {
            key = wait_for_char(keyboard, data_ptr-
>probing_time);
```

```
            if (key != -1) return(key); /* user-terminated */
         }

         pH[index] = pH_probe(meter);

         if (!i--)
         {
            type("\n\t\t");
            i = 5-1;      /* 5 items per line */
         }
         type("%6.3f  ", pH[index]);
      }
      for (index = 0;
        (stddev = deviation(pH, number)) > data_ptr-
>tolerance &&
        (data_ptr->time_limit == UNASSIGNED ||
         get_time() - start_time < data_ptr->time_limit);
        index = ++index % number)
      {
         key = wait_for_char(keyboard, data_ptr-
>probing_time);
         if (key != -1) return(key);     /* user-terminated
*/

         pH[index] = pH_probe(meter);

         if (!i--)
         {
            type("\n\t\t");
            i = 5-1;      /* 5 items per line */
         }
         type("%6.3f  ", pH[index]);
      }
      for (index = 0, sum = 0.0; index < number; ++index)
      {
         sum += pH[index];
      }
      *pH_ptr = sum / number;
      type("%*s: (%.5f+/-%.5f)\n", i * 8, "", *pH_ptr,
stddev);

      return(-1);
}


double
deviation(farray, size)
   float *farray;
   int size;
{
      register int index;
      register double element;
      double sqrt(), sum = 0.0, sum_of_squares = 0.0;
```

```
        for (index = 0; index < size; ++index)
        {
           element = farray[index];
           sum_of_squares += element * element;
           sum += element;
        }
        return(sqrt((size * sum_of_squares - sum * sum) / (size
* (size - 1))));
}
```

Appendix B - Charge Versus pH Transformation Assuming Ideal Electrode

```
#include <stdio.h>
#include <math.h>

     main()
{
      double x, pH, pHo, d, e;
      float a, Np, V, numax, vb, Vo, null, np, nu, b, cb;
      int lines, i;
      char file_in[30], file_out[30];

      FILE *fopen(), *fp;
      FILE *fopen(), *gp;

      printf("name of titration file to be converted ?");
      scanf("%s", file_in);
      printf("name of output file ?");
      scanf("%s", file_out);
      printf("how many lines are there ?");
      scanf("%d",&lines);
      printf("what is the initial volume of protein solution ?");
      scanf("%f",&Vo);
      printf("what is protein conc. mol ?");
      scanf("%f",&np);
      printf("enter max # protons/protein");
      scanf("%f",&numax);
      printf("enter base concentration in mol");
      scanf("%f",&cb);

      fp = fopen(file_in,"r");
      gp = fopen(file_out,"w");

      Vo = Vo/1000;
      Np = Vo * np;

      fscanf(fp,"%f %f",&null,&pHo);
       x = 10.0;

       lines--;
       d = pHo - 14.0;
       a = pow(x,-pHo)-pow(x,d);
       for(i=0; i<lines; i++)
         {
           fscanf(fp,"%f %f", &vb,&pH);
```

```
            e = pH - 14;
            vb = vb/1000000;
            V = Vo + vb;

            b = vb * cb;

 nu=numax+(Vo/Np)*(pow(x,-pHo)-pow(x,d))+(V/Np)*(pow(x,e)-
pow(x,-pH))-b/Np;

            fprintf(gp,"%f %f\n", pH, nu);
          }
    }
```

Appendix C - Acidic Range Charge Versus pH Transformation Including Water Corrections.

```
#include <stdio.h>
#include <math.h>

     main()
{
      double
x,pH,pHo,d,e,pHw1o,aw,dvw1,dvw2,err1,f1,f2,pHw2o,pHw,nu1,nu2;
      double
bw,dvw1a,dvw2a,dvw1b,dvw2b,pHw1a,pHw1b,pHw2a,pHw2b,f,check1,c
heck2;
      double
eee1,eee2,dacid,dVo,dconc,erra,errv,errc,error,fdacid,fdVo,fd
conc;
      double sqerror,half;
      float a, Np, V, numax, vb, Vo, null, np, nu, b, cb;
      char
file_in[30],file_out[30],water1[30],water2[30],log[30],salt[3
0];
      char date[15],com[100];

      FILE *fopen(), *fp, *gp, *wp1, *wp2, *lp;

      printf("name of titration file to be converted? ");
      scanf("%s", file_in);
      printf("name of first blank water curve to correct dV
vs pH? ");
      scanf("%s", water1);
      printf("name of second blank water curve to correct dV
vs pH? ");
      scanf("%s", water2);
      printf("name of output file? ");
      scanf("%s", file_out);
      printf("what is the initial volume of protein solution?
");
      scanf("%f",&Vo);
      printf("what is protein conc. mol? ");
      scanf("%f",&np);
      printf("enter initial # protons/protein: ");
      scanf("%f",&numax);
      printf("enter acid concentration in mol: ");
      scanf("%f",&cb);
      printf("salt conditions?");
      scanf("%s",salt);
```

```
      printf("experiment date?");
      scanf("%s",date);
      printf("comment?-- use _ to connect words(<100
characters)");
      scanf("%s",com);

      printf("enter fractional uncertainty in acid conc.: ");
      scanf("%f",&fd  id);
      printf("enter .. ctional uncertainty in initial volume:
");
      scanf("%f",&fdVo);
      printf("enter fractional uncertainty in initial protein
conc.: ");
      scanf("%f",&fdconc);

      sprintf(log,"log%s",file_out);
      lp = fopen(log,"w");

      fprintf(lp,"exp. date: %s\n",date);
      fprintf(lp,"salt conditions = %s\n",salt);
      fprintf(lp,"acid conc. = %fM\n",cb);
      fprintf(lp,"protein conc. = %fM\n",np);
      fprintf(lp,"volume = %fml\n",Vo);
      fprintf(lp,"fractional uncertainty in acid conc. =
%f\n",fdacid);
      fprintf(lp,"fractional uncertainty in acid conc. =
%f\n",fdVo);
      fprintf(lp,"fractional uncertainty in protein conc. =
%f\n",fdconc);
      fprintf(lp,"comments: %s\n",com);

      fp = fopen(file_in,"r");
      gp = fopen(file_out,"w");

      Vo = Vo/1000;
      Np = Vo * np;

      fscanf(fp,"%f %f",&null,&pHo);
       x = 10.0;

       printf("\n.");
       d = pHo - 14.0;
       a = pow(x,-pHo)-pow(x,d);

     while(fscanf(fp, "%f %f", &vb,&pH) == 2) /* while not
end of file */
        {
       printf(".");
       wp1 = fopen(water1,"r");
       wp2 = fopen(water2,"r");

       fscanf(wp1,"%f %f",&null,&pHw1o);
       fscanf(wp2,"%f %f",&null,&pHw2o);
```

```
          do{
                    fscanf(wp1,"%f %f",&dvw1a,&pHw1a);
              if (fscanf(wp1,"%f %f",&dvw1b,&pHw1b) == EOF)
              {
                 printf("\nEnd of 1st water file!\n");
                 exit(1);
              }
              aw = pHw1b - pH;
            }while(aw > 0);

          do{
                fscanf(wp2,"%f %f",&dvw2a,&pHw2a);
             if (fscanf(wp2,"%f %f",&dvw2b,&pHw2b) == EOF)
             {
                   printf("\nEnd of 2nd water file!\n");
                   exit(1);
             }
             bw = pHw2b - pH;
              }while(bw > 0);


            fclose(wp1);
            fclose(wp2);

          dvw1a = dvw1a/1000000;
          dvw2a = dvw2a/1000000;
          dvw1b = dvw1b/1000000;
          dvw2b = dvw2b/1000000;

dvw1=(pow(x,-pHw1a)-pow(x,-pH))*((dvw1b-dvw1a)/(pow(x,-
pHw1a)-pow(x,-pHw1b)));
 dvw1 = dvw1+dvw1a;

dvw2=(pow(x,-pHw2a)-pow(x,-pH))*((dvw2b-dvw2a)/(pow(x,-
pHw2a)-pow(x,-pHw2b)));
 dvw2 = dvw2+dvw2a;
             e = pH - 14;
             vb = vb/1000000;
             V = Vo + vb;
             b = vb * cb;
          eee1 = pHw1o - 14;
          eee2 = pHw2o - 14;

f1=(Vo+dvw1)*(pow(x,-pH)-pow(x,e))/(Vo*(pow(x,-pHw1o)-
pow(x,eee1))+cb*dvw1);
f2=(Vo+dvw2)*(pow(x,-pH)-pow(x,e))/(Vo*(pow(x,-pHw2o)-
pow(x,eee2))+cb*dvw2);

          /*    check1 = pH-pHw1o;
          check2 = pH-pHw2o;

         if(check1>0)
            f1=1;
          if(check2>0)
```

```
          f2=1;  */

        f = (f1+f2)/2;

 nu=numax+(Vo/Np)*(pow(x,-pHo)-pow(x,d))+(V/Np)*(pow(x,e)-
pow(x,-pH))/f+b/Np;

nu1=numax+(Vo/Np)*(pow(x,-pHo)-pow(x,d))+(V/Np)*(pow(x,e)-
pow(x,-pH))/f1+b/Np;
nu2=numax+(Vo/Np)*(pow(x,-pHo)-pow(x,d))+(V/Np)*(pow(x,e)-
pow(x,-pH))/f2+b/Np;

err1 = (nu1 - nu2)/2;

dacid = fdacid*cb;
dVo = fdVo*Vo;
dconc = fdconc*Np/Vo;

erra = (vb/Np)*dacid;
errv = ((dVo*vb)/(Vo*Np))*((pow(x,e)-pow(x,-pH))/f+cb);
errc = dconc*Vo*(nu-numax)/Np;

sqerror = err1*err1+erra*erra+errv*errv+errc*errc;
half = .5;

error = pow(sqerror,half);

/*   if(err1<0)
 err1 = -err1;  */

           fprintf(gp,"%f %f ?%f\n", pH, nu, error);
         }
      printf("\n");
    }
```

Appendix D - Basic Range Charge Versus pH Transformation Including Water Corrections.

```
#include <stdio.h>
#include <math.h>

      main()
{
       double
x,pH,pHo,d,e,pHw1o,aw,dvw1,dvw2,err1,f1,f2,pHw2o,pHw,nu1,nu2;
      double
bw,dvw1a,dvw2a,dvw1b,dvw2b,pHw1a,pHw1b,pHw2a,pHw2b,f,check1,check2;
       double eee1,eee2,xxx1,xxx2,yyy1,yyy2,zzz1,zzz2;
 double
dbase,dVo,dconc,errb,errv,errc,error,fdbase,fdVo,fdconc,sqerror,half
;
       float a, Np, V, numax, vb, Vo, null, np, nu, b, cb;
       char
file_in[30],file_out[30],water1[30],water2[30],salt[30],log[30];
       char date[15],com[100];

       FILE *fopen(), *fp, *gp, *wp1, *wp2, *lp;

       printf("name of titration file to be converted? ");
       scanf("%s", file_in);
       printf("name of first blank water curve to correct dV vs pH?
");
       scanf("%s", water1);
       printf("name of second blank water curve to correct dV vs pH?
");
       scanf("%s", water2);
       printf("name of output file? ");
       scanf("%s", file_out);
       printf("what is the initial volume of protein solution? ");
       scanf("%f",&Vo);
       printf("what is protein conc. mol? ");
       scanf("%f",&np);
       printf("enter max # protons/protein: ");
       scanf("%f",&numax);
       printf("enter base concentration in mol: ");
       scanf("%f",&cb);
       printf("salt conditions?");
       scanf("%s",salt);
       printf("experiment date?");
       scanf("%s",date);
       printf("comments?--<100 chars., connect words with _");
```

```
      scanf("%s",com);


      printf("enter fractional uncertainty in base conc.: ");
      scanf("%f",&fdbase);
      printf("enter fractional uncertainty in initial volume: ");
      scanf("%f",&fdVo);
      printf("enter fractional uncertainty in protein conc.: ");
      scanf("%f",&fdconc);


      sprintf(log,"log%s",file_out);
      lp = fopen(log,"w");

      fprintf(lp,"exp. date: %s\n",date);
      fprintf(lp,"salt conditions = %s\n",salt);
fprintf(lp,"base conc. = %fM\n",cb);
      fprintf(lp,"protein conc. = %fM\n",np);
      fprintf(lp,"volume = %fml\n",Vo);
      fprintf(lp,"fractional uncertainty in base conc. =
%f\n",fdbase);
      fprintf(lp,"fractional uncertainty in protein conc. =
%f\n",fdconc);
      fprintf(lp,"fractional uncertainty in volume = %f\n",fdVo);
      fprintf(lp,"comments: %s\n",com);

      fp = fopen(file_in,"r");
      gp = fopen(file_out,"w");

      Vo = Vo/1000;
      Np = Vo * np;

      fscanf(fp,"%f %f",&null,&pHo);
       x = 10.0;

       printf("\n.");
       d = pHo - 14.0;
       a = pow(x,-pHo)-pow(x,d);

     while(fscanf(fp, "%f %f", &vb,&pH) == 2) /* while not end of
file */
        {
       printf(".");
       wp1 = fopen(water1,"r");
       wp2 = fopen(water2,"r");

       fscanf(wp1,"%f %f",&null,&pHw1o);
       fscanf(wp2,"%f %f",&null,&pHw2o);

       do{
                fscanf(wp1,"%f %f",&dvw1a,&pHw1a);
           if (fscanf(wp1,"%f %f",&dvw1b,&pHw1b) == EOF)
           {
              printf("\nEnd of 1st water file!\n");
```

```
                exit(1);
              }
              aw = pHw1b - pH;
            }while(aw < 0);

          do{
                fscanf(wp2,"%f %f",&dvw2a,&pHw2a);
             if (fscanf(wp2,"%f %f",&dvw2b,&pHw2b) == EOF)
             {
                   printf("\nEnd of 2nd water file!\n");
                   exit(1);
             }
             bw = pHw2b - pH;
              }while(bw < 0);


            fclose(wp1);
            fclose(wp2);

          dvw1a = dvw1a/1000000;
          dvw2a = dvw2a/1000000;
          dvw1b = dvw1b/1000000;
          dvw2b = dvw2b/1000000;

dvw1=(pow(x,pHw1a)-pow(x,pH))*((dvw1b-dvw1a)/(pow(x,pHw1a)-
pow(x,pHw1b)));
dvw1 = dvw1+dvw1a;

dvw2=(pow(x,pHw2a)-pow(x,pH))*((dvw2b-dvw2a)/(pow(x,pHw2a)-
pow(x,pHw2b)));
dvw2 = dvw2+dvw2a;

             e = pH - 14;
             vb = vb/1000000;
             V = Vo + vb;
             b = vb * cb;
          eee1 = pHw1o-14;
          eee2 = pHw2o-14;

          yyy1 = cb*dvw1;
             yyy2 = cb*dvw2;
          zzz1 = -Vo*(pow(x,-pHw1o)-pow(x,eee1));
          zzz2 = -Vo*(pow(x,-pHw2o)-pow(x,eee2));

          f1 = (Vo+dvw1)*(pow(x,e)-pow(x,-pH))/(yyy1+zzz1);
          f2 = (Vo+dvw2)*(pow(x,e)-pow(x,-pH))/(yyy2+zzz2);

          check1 = pH-pHw1o;
          check2 = pH-pHw2o;

      /*    if(check1>0)
            f1=1;
          if(check2>0)
            f2=1;   */
```

```
        f = (f1+f2)/2;

 nu=numax+(Vo/Np)*(pow(x,-pHo)-pow(x,d))+(V/Np)*(pow(x,e)-pow(x,-
pH))/f-b/Np;

 /*  printf("%f %f %f\n",f1,f2,f);  */

nu1=numax+(Vo/Np)*(pow(x,-pHo)-pow(x,d))+(V/Np)*(pow(x,e)-pow(x,-
pH))/f1-b/Np;
nu2=numax+(Vo/Np)*(pow(x,-pHo)-pow(x,d))+(V/Np)*(pow(x,e)-pow(x,-
pH))/f2-b/Np;

err1 = (nu1 - nu2)/2;

dbase = fdbase*cb;
dVo = fdVo*Vo;
dconc = fdconc*Np/Vo;

errb = (vb/Np)*dbase;
errv = (dVo*vb/(Vo*Np))*((pow(x,e)-pow(x,-pH))/f-cb);
errc = dconc*(nu-numax)*Vo/Np;

sqerror = err1*err1+errb*errb+errv*errv+errc*errc;
half = .5;

error = pow(sqerror,half);

           fprintf(gp,"%f %f ?%f\n", pH, nu, error);
         }
      printf("\n");
    }
```

Appendix E - Linderstrom-Lang Titration Curve Software

```
#include <stdio.h>
#include <math.h>

main()
{

     double
w,Z,pH,deltpH,deltZ,err1,err2,ten,Zbest,func,f1,f2,f3,f4,f5,f
6;
     int m,n,mmax,nmax;
     char graph_out[30];

     FILE *fopen(),*gp;

/*   w = .069;*/
     w=0;
     ten = 10.0;
     nmax = 100;
     mmax = 200;
     pH = 2;
     err2 = 100000000;

     gp = fopen(graph_out,"w");

     printf("graph out");
     scanf("%s",graph_out);

     for(n=1;n<=nmax;n++)
          {

          Z = -30;

          for(m=1;m<=mmax;m++)
               {
               func = 0;
               f1=pH-4+w*Z;
               func=13/(1+pow(ten,f1));
               f2=pH-4.5+w*Z;
               func = func+9/(1+pow(ten,f2));
               f3=pH-6.4+w*Z;
               func=func+5/(1+pow(ten,f3));
               f4=pH-10+w*Z;
```

```
        func=func+1/(1+pow(ten,f4));
        f5=pH-10.4+w*Z;
        func=func+3/(1+pow(ten,f5));
        f6=pH-12+w*Z;
        func=func+20/(1+pow(ten,f6));

        func=func-23;

        err1=(Z-func)*(Z-func);

        if(err2>err1){
             Zbest=Z;
             err2=err1;
             }else{
                  err2=err1;
                  }

        deltZ = 0.333;
        Z=Z+deltZ;
        }

        fprintf(gp,"%f %f\n",pH,Zbest);
        printf("%f %f\n",pH,Zbest);

        deltpH=0.1;
        pH=pH+deltpH;
        }

fclose(gp);

}
```

Appendix F - Kirkwood-Tanford Titration Curve Software

```
#include <stdio.h>
#include <math.h>


main()
{

double Wr[150][150],d[150][8],pHpK[200][150];
double a,b,c,e,rr,pH,nu,x,ten,pH2,y,rrr;
int i,j,k,ll,l,nn,num,m,n,jjj,jj,kkk,jjmax,kk,mm,nmax,jmax;
char file_in[30],file_out[30],graph_out[30];

FILE *fopen(),*gp;
FILE *fopen(),*lp;
FILE *fopen(),*fp;

a = 22.67;
b = .12899;
c = .03;
e = 2.7183;
ten = 10.0;

printf("enter the name of the input file (sigma pKo x y z
sa)");
scanf("%s",file_in);

printf("enter the name of the pK out file");
scanf("%s",file_out);

printf("enter the name of the graph out file");
scanf("%s",graph_out);

printf("enter number of iterations");
scanf("%d",&num);

fp = fopen(file_in,"r");
gp = fopen(graph_out,"w");
lp = fopen(file_out,"w");

fscanf(fp,"%d",&nmax);
for(n=1;n<=nmax;n++)
  {
```

```
fscanf(fp,"%f %f %f %f
%f",&d[n][1],&d[n][2],&d[n][3],&d[n][4],&d[n][5]);
fscanf(fp,"%f",&d[n][8]);
  }

for(m=1;m<=nmax;m++)
  {
    for(l=1;l<=nmax;l++)
      {
      rr = (d[l][3]-d[m][3])*(d[l][3]-d[m][3]);
      rr = rr + (d[l][4]-d[m][4])*(d[l][4]-d[m][4]);
      rr = rr + (d[l][5]-d[m][5])*(d[l][5]-d[m][5]);
         rr = sqrt(rr);
      rrr = -b*rr;

      Wr[l][m] = 236.5*a*pow(e,rrr)/((rr + c)*1000);
      Wr[l][m] = (1-(d[l][8]+d[m][8])/2)*Wr[l][m];

        }
  }

jmax = 200;
for(j=1;j<=jmax;j++)
  {

    pH = 1.95 + j*(0.05);

for(ll=1;ll<=nmax;ll++)
  {
    d[ll][6] = d[ll][2];
  }

printf(".");
    for(nn=1;nn<=num;nn++)
      {
     for(i=1;i<=nmax;i++)
       {
         d[i][7] = 0;

         for(k=1;k<=nmax;k++)
           {
     d[i][7] = d[i][7]-(1/(1+pow(ten,pH-d[k][6]))+(d[k][1]-
1)/2)*(Wr[i][k]);
           }

/*d[i][6]=
d[i][7]+d[i][2]+(1/(1+pow(ten,pH-d[i][6]))-(d[i][1]-
1)/2)*(Wr[i][i]);*/

d[i][6]=d[i][7]+d[i][2]+(1/(1+pow(ten,pH-d[i][6]))+(d[i][1]-
1)/2)*(Wr[i][i]);
       }
        for(kk=1;kk<=nmax;kk++)
          {
```

```
                pHpK[j][kk] = d[kk][6];

            }
        }
      nu = 0;

    for(mm=1;mm<=nmax;mm++)
      {
        x = pH - pHpK[j][mm];
        nu = nu + 1/(1+pow(ten,x));
      }
    fprintf(gp,"%f %f\n",pH,nu);
    }
jjmax = 10;

for(jj=1;jj<=jjmax;jj++)
   {
      jjj = 20*jj;

    pH2 = 1.0 + jjj*(0.05);

     for(kkk=1;kkk<=nmax;kkk++)
      {
     y = pHpK[jjj][kkk];
        fprintf(lp,"%f %f\n",pH2,y);
      }
   }
fclose(lp);
fclose(gp);
fclose(fp);
}
```

Appendix G - Enlarged Charge Versus pH Curves for the γ-Crystallins

This appendix contains enlargements of figures 4.7 through 4.10. These figures are reproduced in enlarged form so that the charge can be more accurately read from the figures. All of the data shown in this section was collected in 0.1 M KCl aqueous solvent.

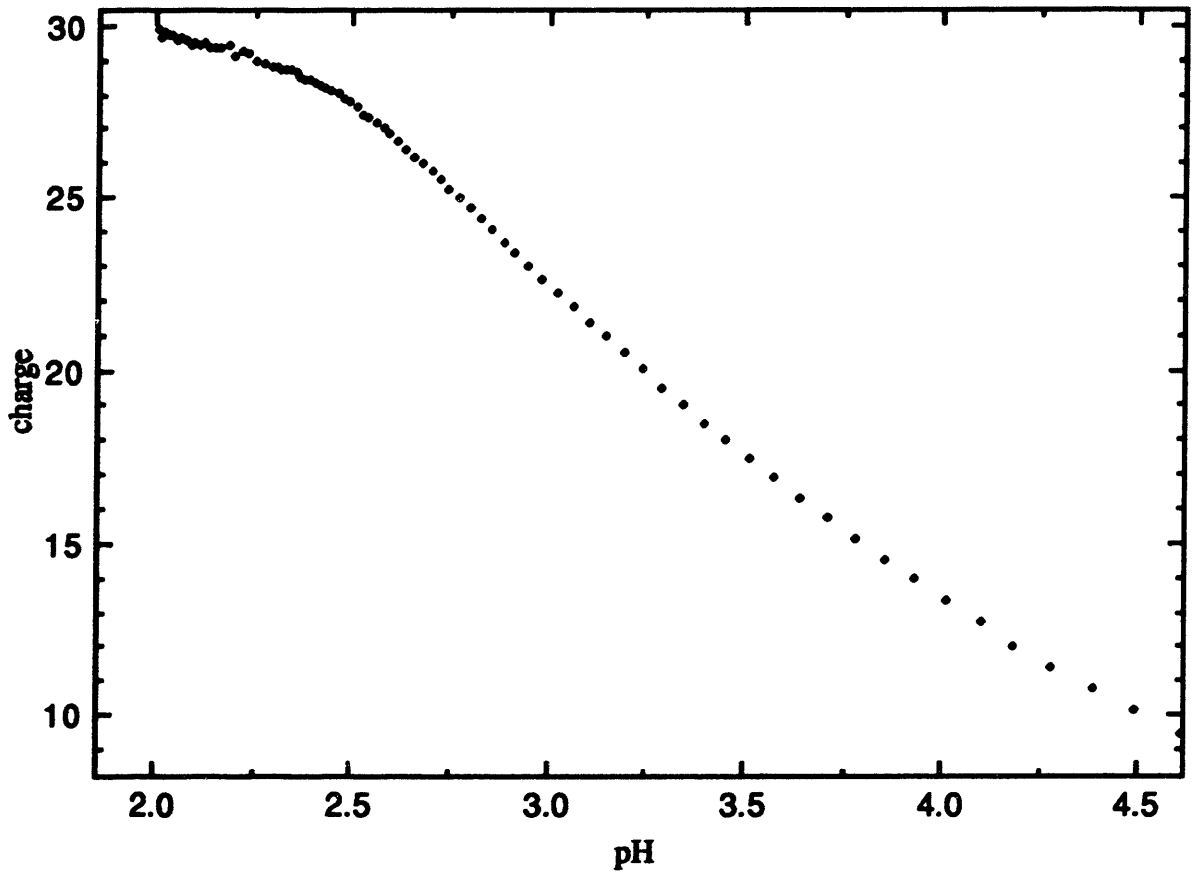


Figure G1 - γ_{II} charge versus pH in 0.1M KCl - acidic range data.

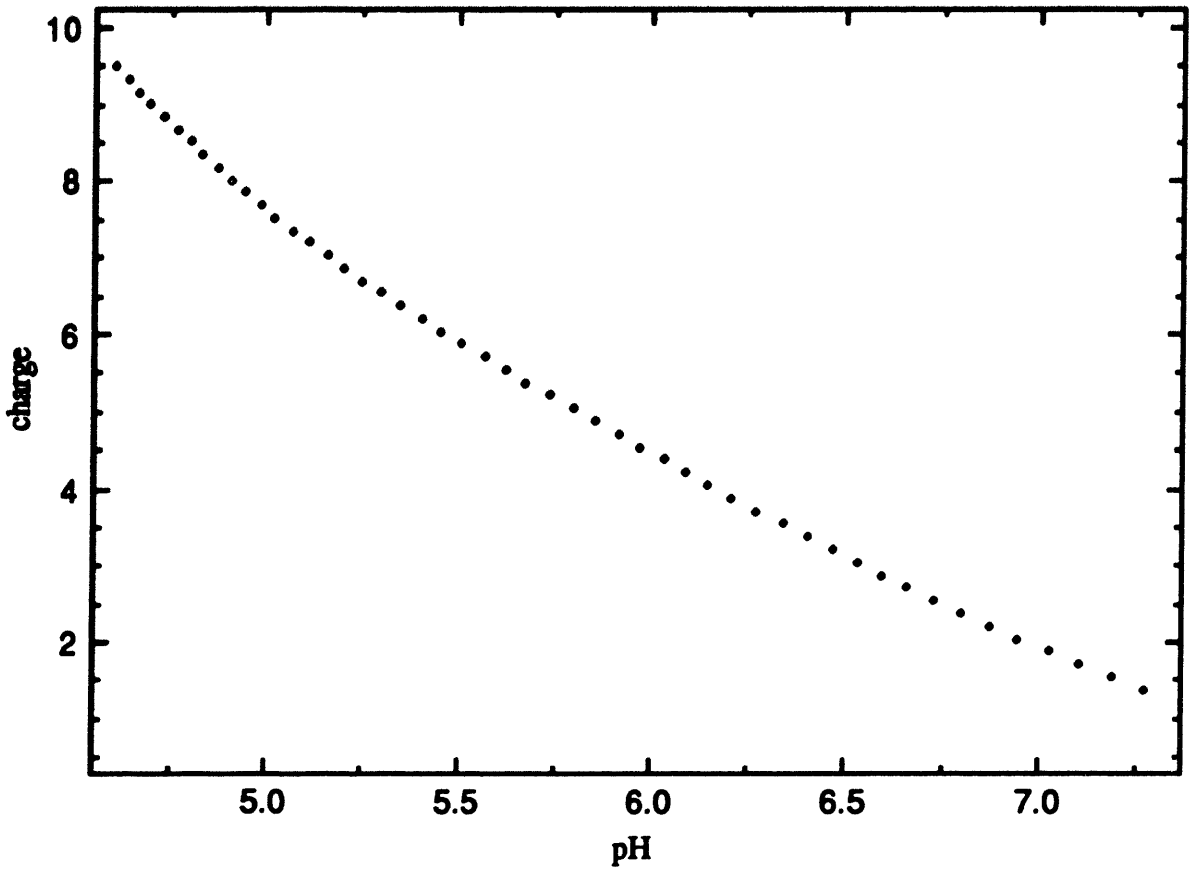


Figure G2 - γ_{II} charge versus pH in 0.1M KCl - middle range data.

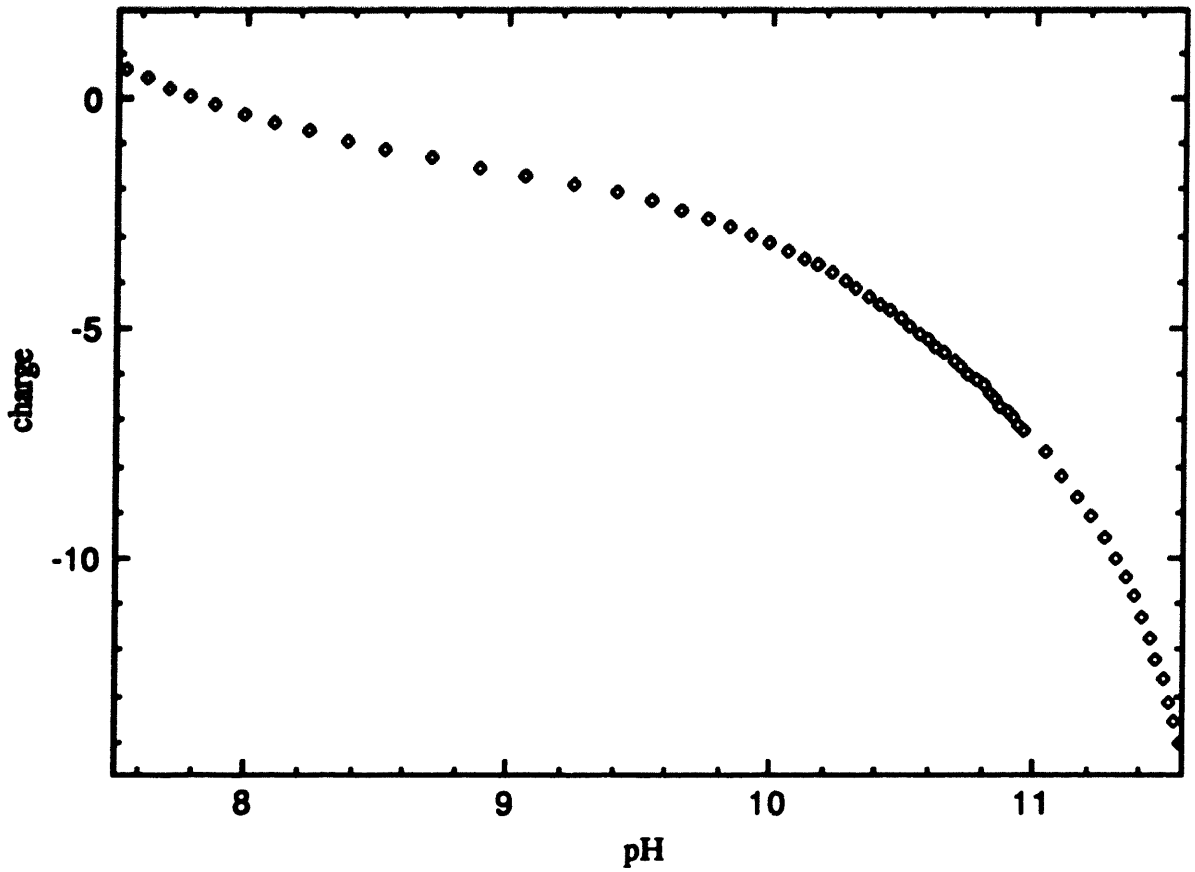


Figure G3 - γ_{II} charge versus pH in 0.1M KCl - basic range data.

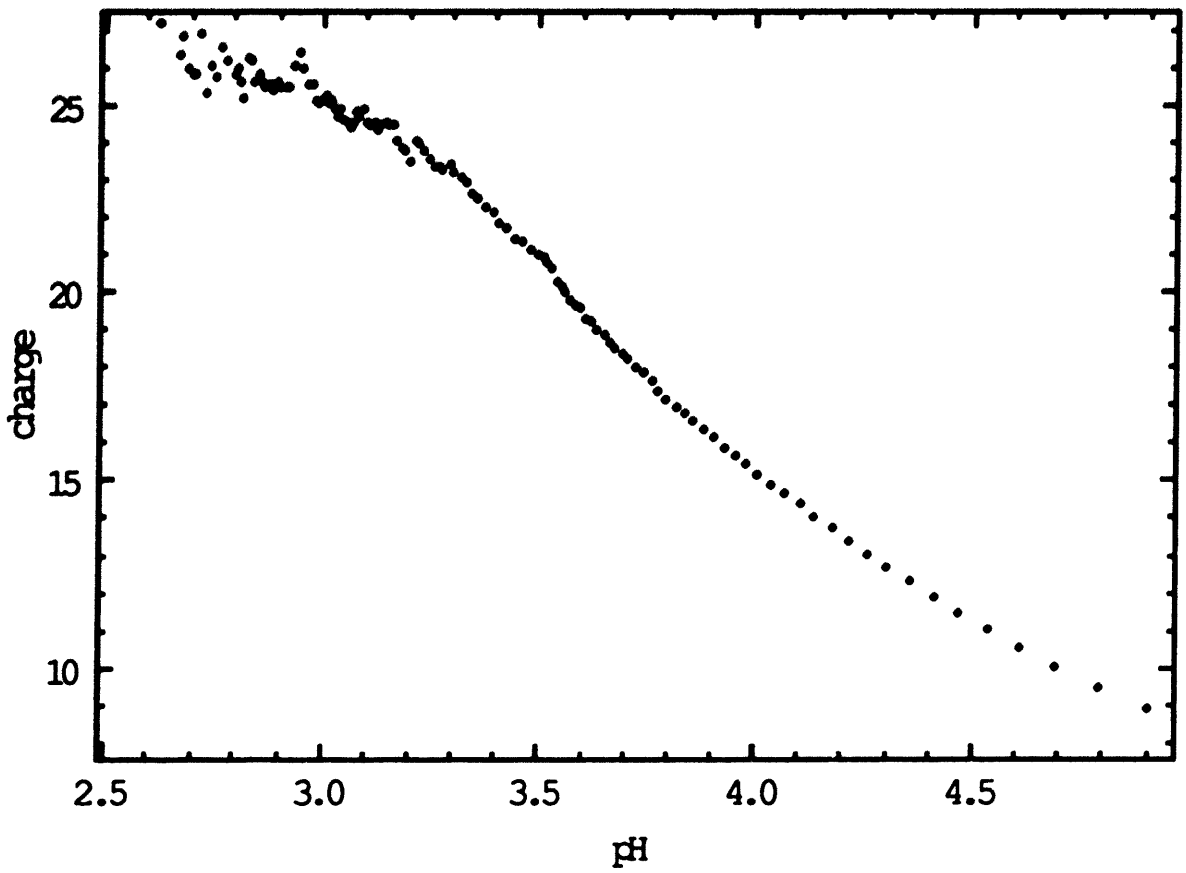


Figure G4 - γ_{IIIa} charge versus pH in 0.1M KCl - acidic range data.

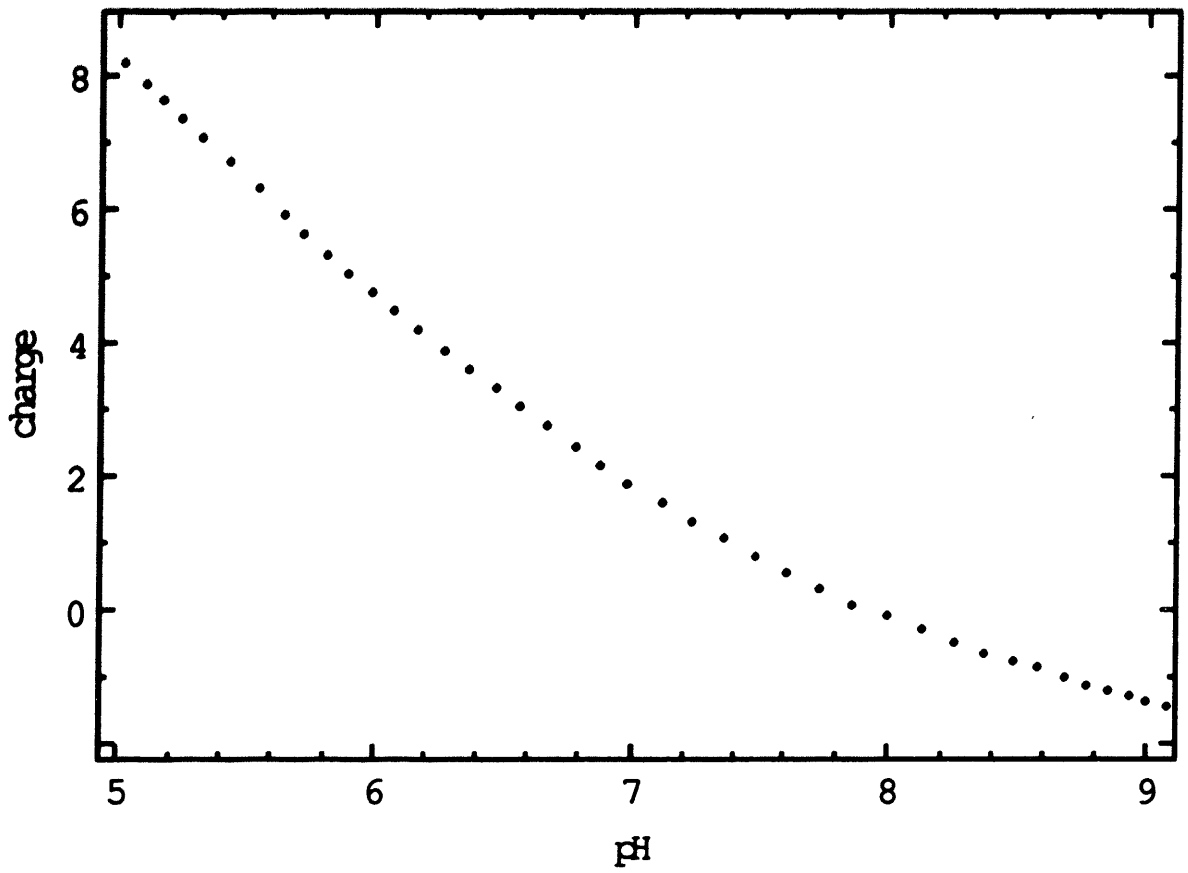


Figure G5 - γ_{IIIa} charge versus pH in 0.1M KCl - middle range data.

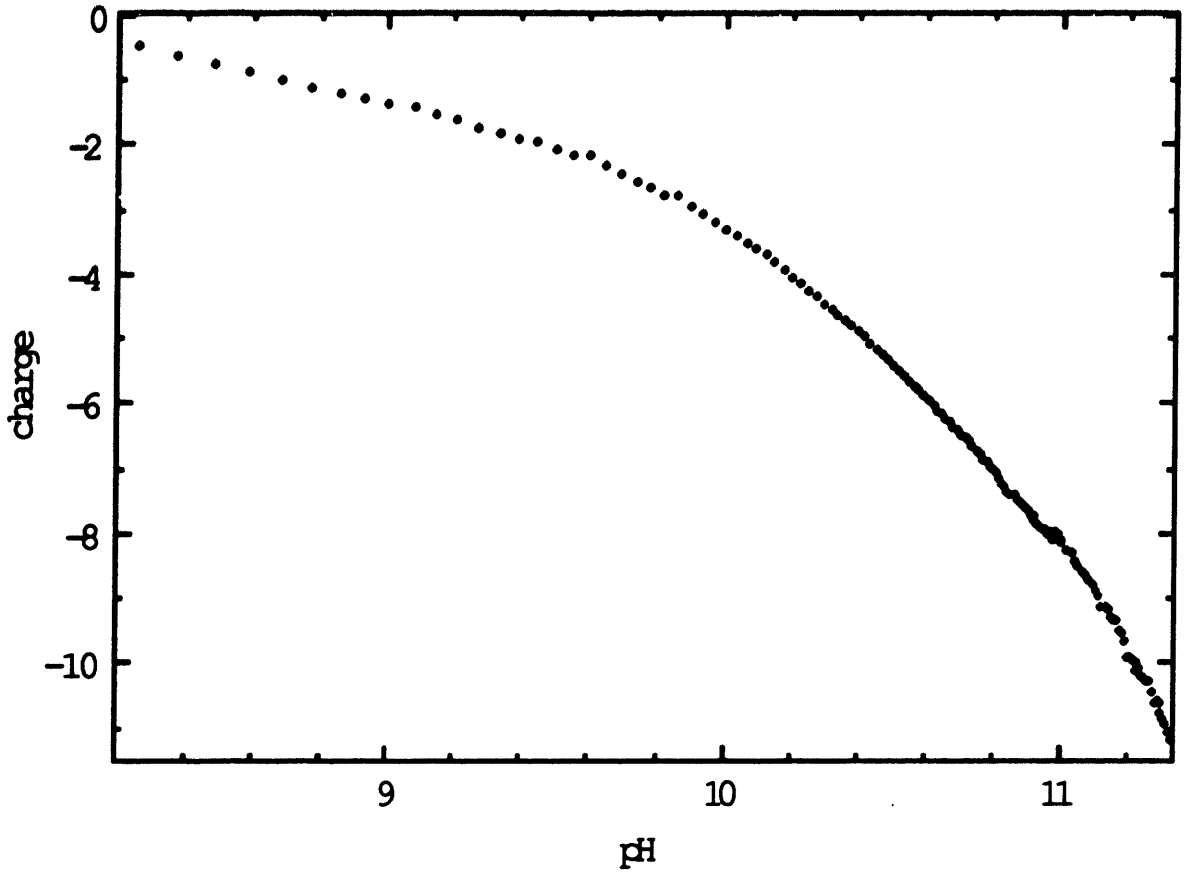


Figure G6 - γ_{IIIa} charge versus pH in 0.1M KCl - basic range data.

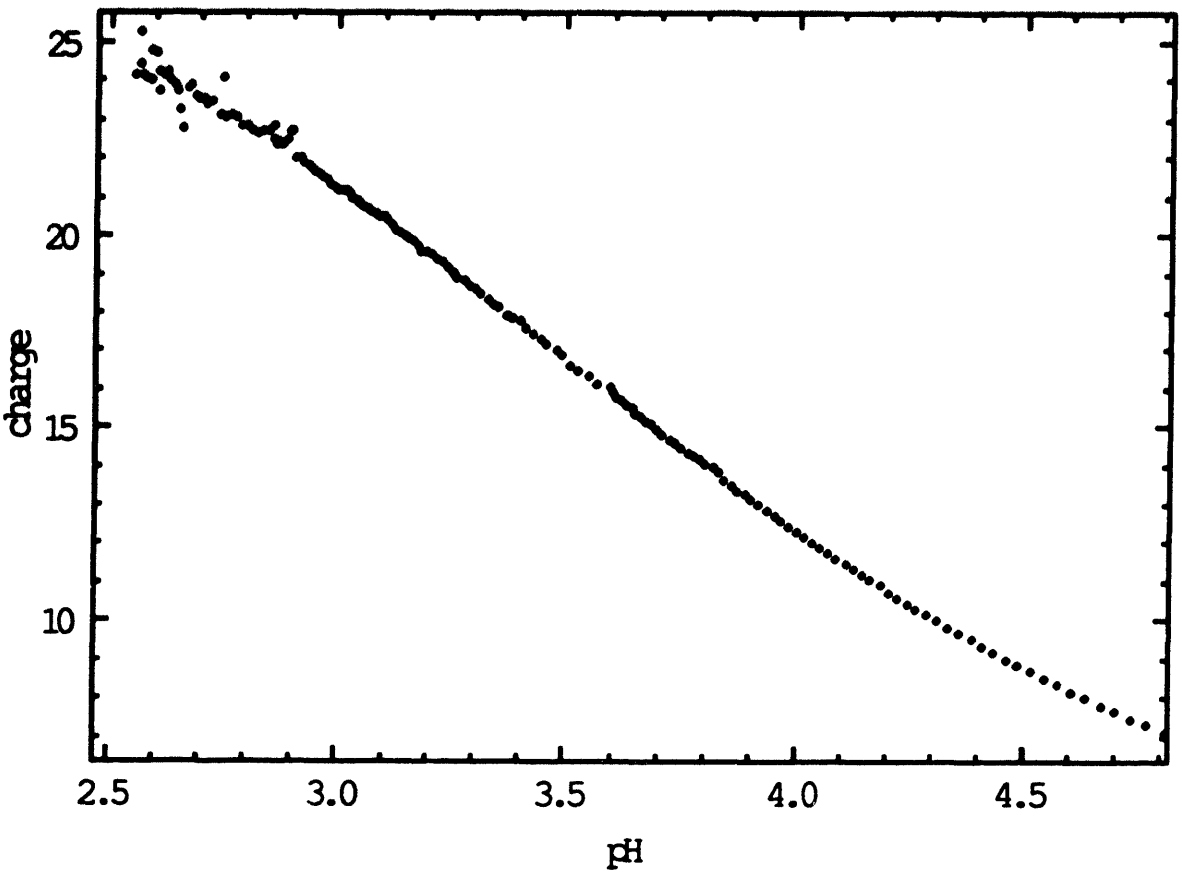


Figure G7 - γ_{IIIb} charge versus pH in 0.1M KCl - acidic range data.

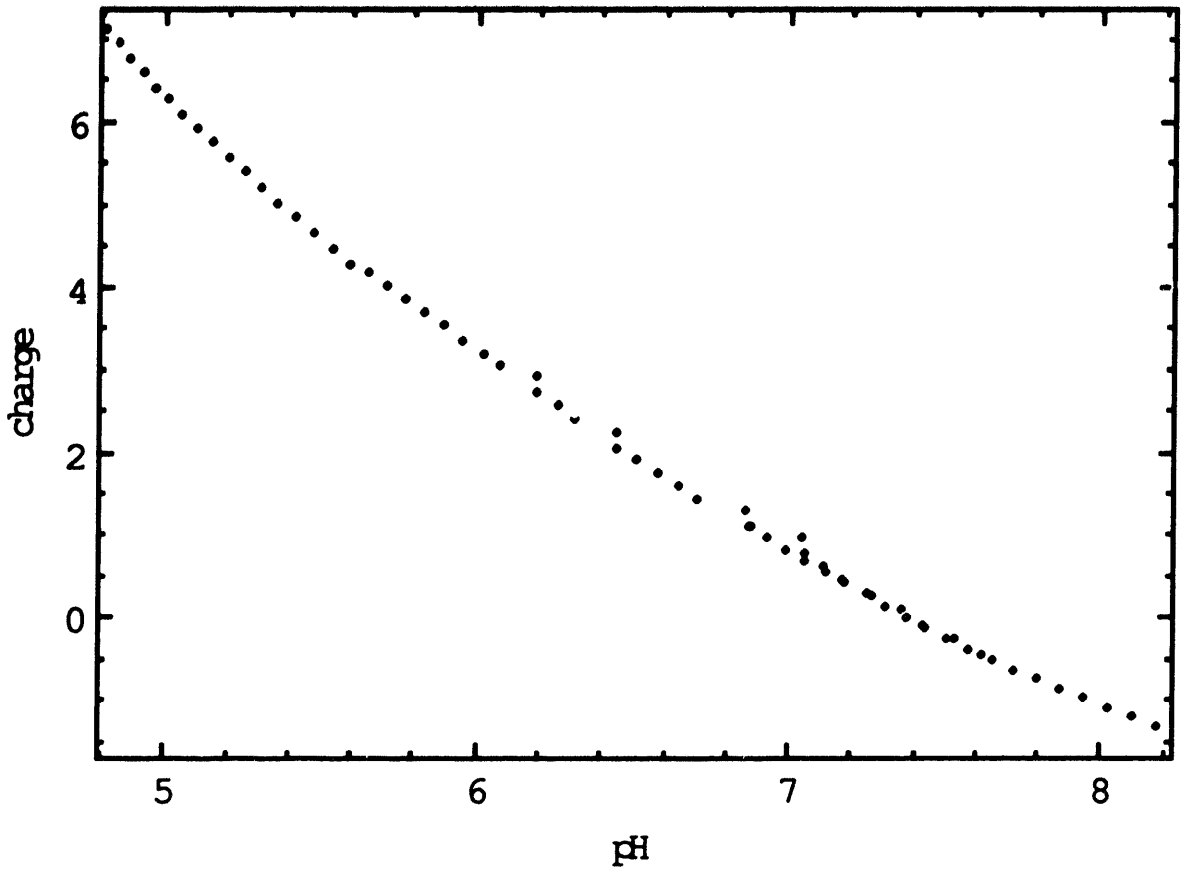


Figure G8 - γ_{IIIb} charge versus pH in 0.1M KCl - middle range data.

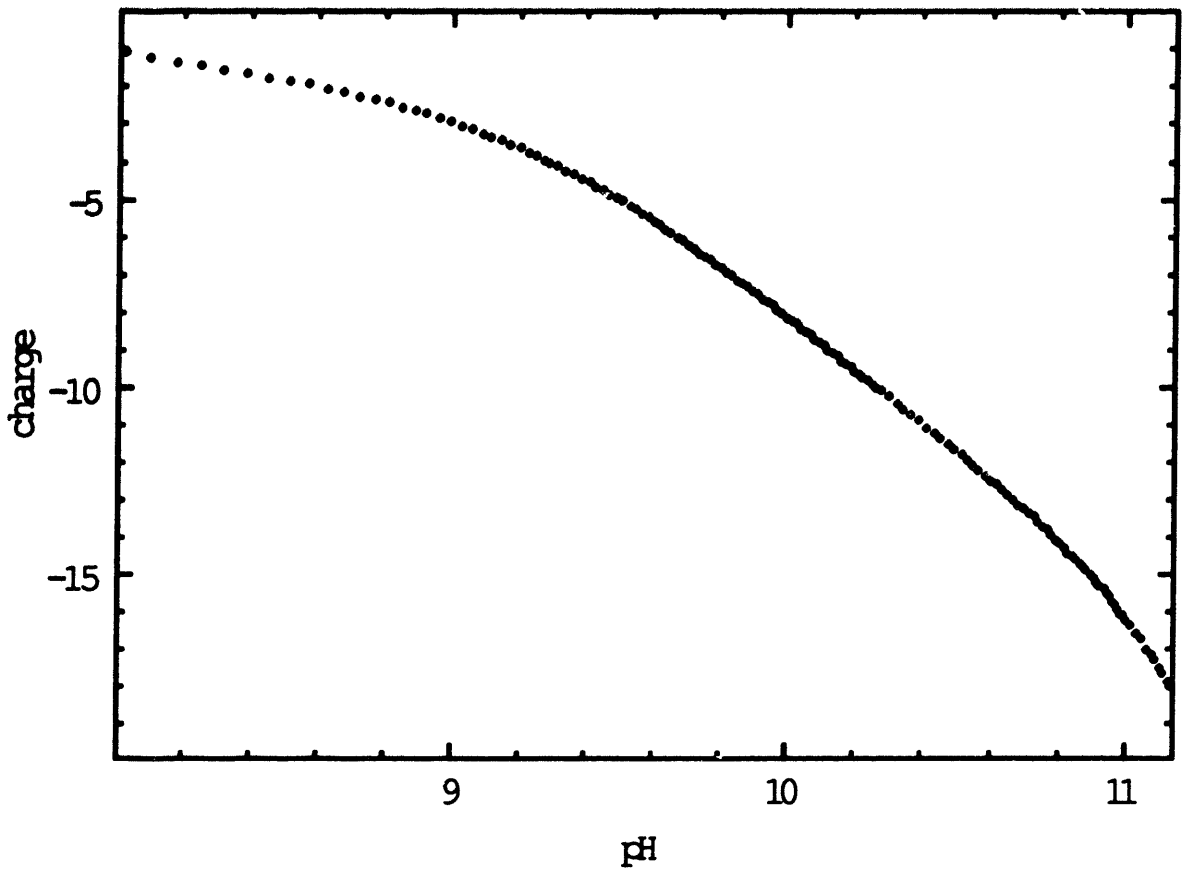


Figure G9 - γ_{IIIb} charge versus pH in 0.1M KCl - basic range data.

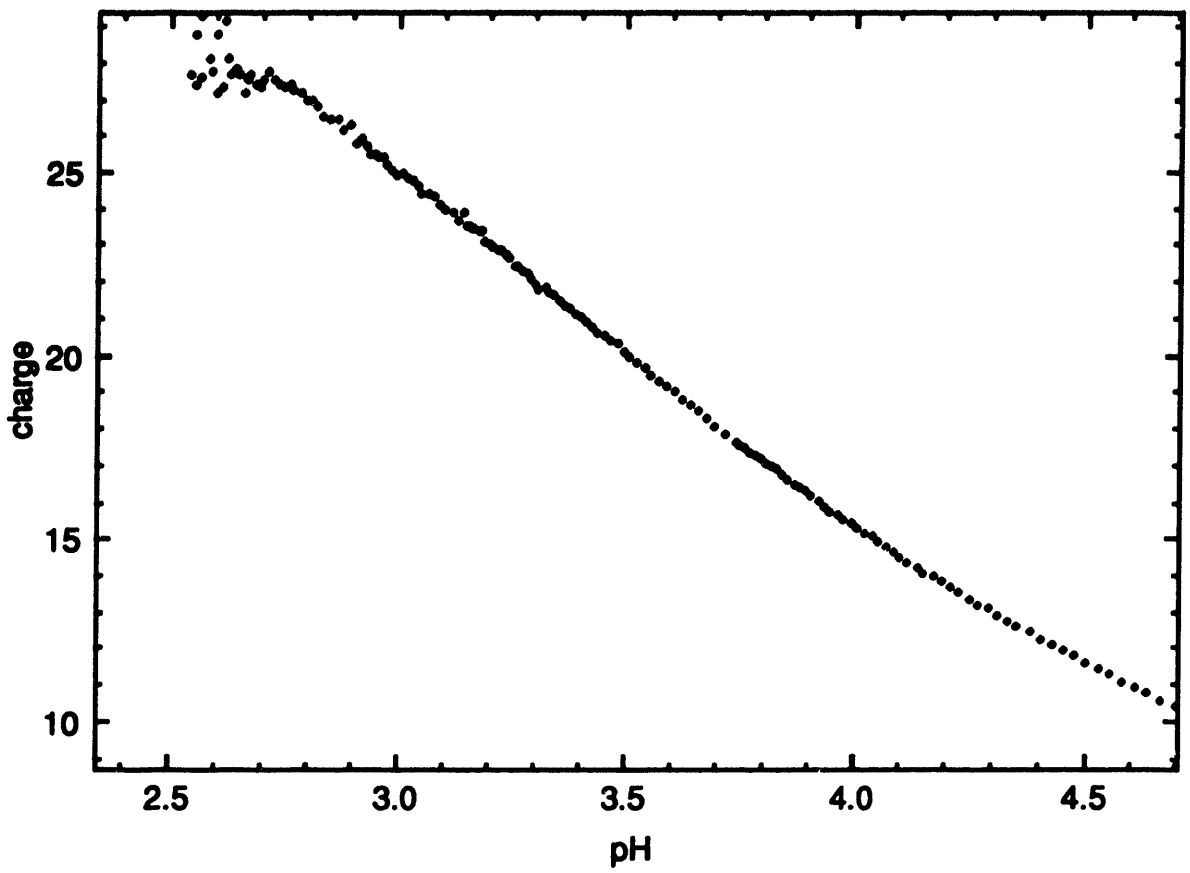


Figure G10 - γ_{IV} charge versus pH in 0.1M KCl - acidic range data.

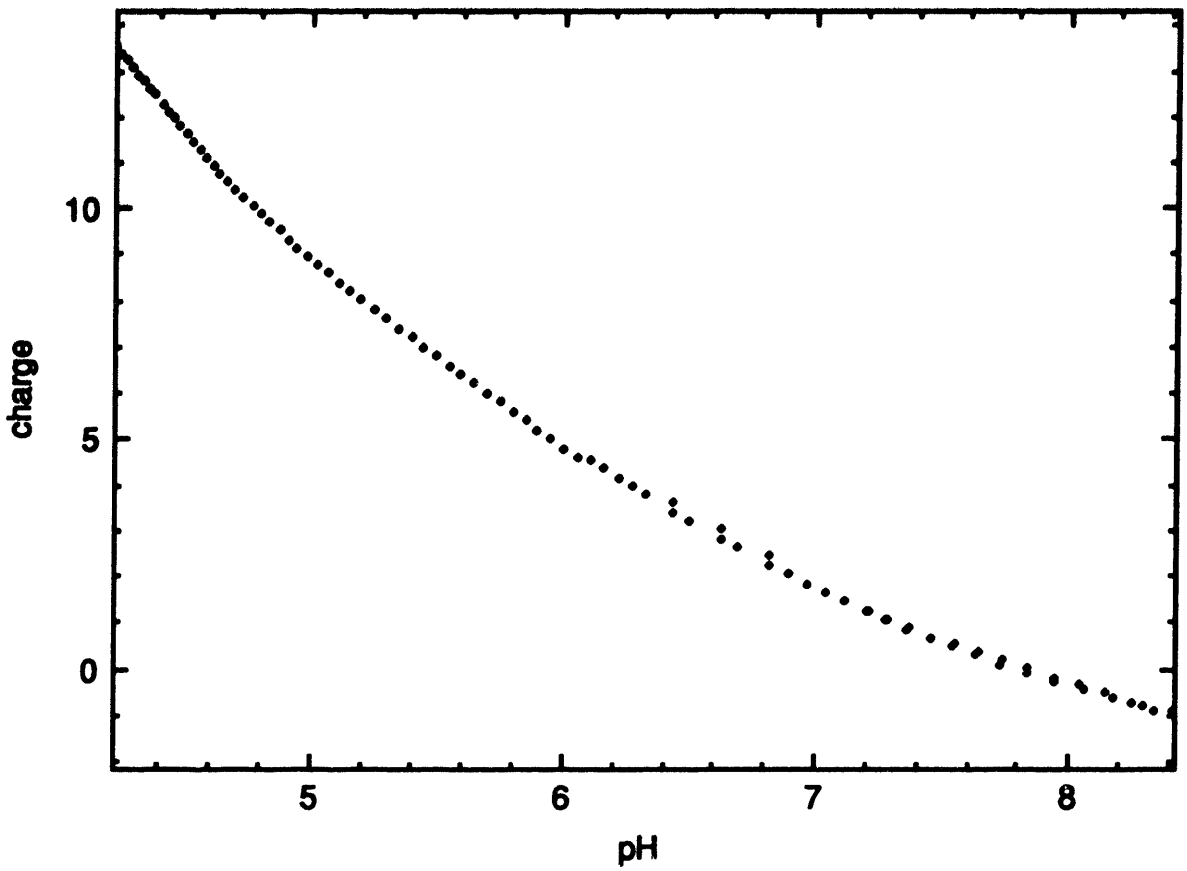


Figure G11 - γ_{IV} charge versus pH in 0.1M KCl - middle range data.

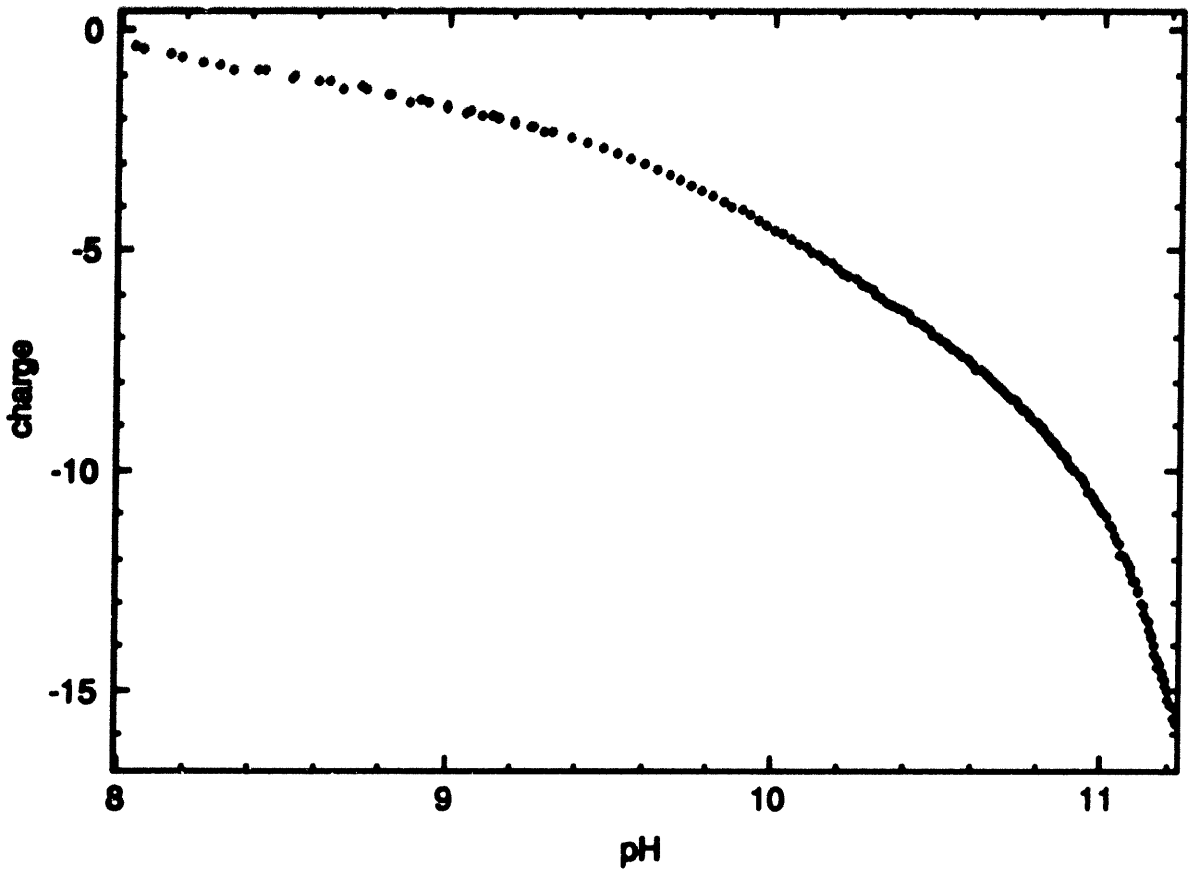


Figure G12 - γ_{IV} charge versus pH in 0.1M KCl - basic range data.